



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 198439

TO: Ginny Portner
Location: REM-3C18
Art Unit: 1645
Friday, August 11, 2006

From: Paul Schulwitz
Location: Biotech-Chem Library
REM-1A65
Phone: 571-272-2527

Case Serial Number: 09/904994

Paul.schulwitz@uspto.gov

Search Notes

Examiner Portner,

Please review the attached search results.

If you have any questions or if you would like to refine the search query, please feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
REM-1A65
571-272-2527

Schulwitz, Paul

From: Portner, Ginny
Sent: Friday, August 04, 2006 11:32 AM
To: Schulwitz, Paul
Subject: 09/904994 nucleic and amino acid sequence search with specific gap values; claims attached
Importance: High

Thanks for your help with this search issue. I am sending the claims to you now (Snagit training WORKed!!!!!!!)
Once I check the Spec for a program name etc I will send you another email. ginny

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02; Mail Box 3C18
(571) 272-0862

Claims:
23
34
40

72798

Schulwitz, Paul

From: Portner, Ginny
Sent: Friday, August 04, 2006 11:44 AM
To: Schulwitz, Paul
Subject: FW: 09/904994 nucleic and amino acid sequence search with specific gap values; claims attached
Importance: High

Paul, From the Specification, I found that they sued a computer program called "Align Plus for Windows, available from Scientific and Educational Software, P.O. Box 72045, Durham, NC 27722-2045 USA. I hope that this information helps. ginny

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02; Mail Box 3C18
(571) 272-0862

-----Original Message-----

From: Portner, Ginny
Sent: Friday, August 04, 2006 11:32 AM
To:Subject: 09/904994 nucleic and amino acid sequence search with specific gap values; claims attached
Importance: High

Thanks for your help with this search issue. I am sending the claims to you now (Snagit training WORKed!!!!!!!)
Once I check the Spec for a program name etc I will send you another email. ginny

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02; Mail Box 3C18
(571) 272-0862

Schulwitz, Paul

From: Portner, Ginny
Sent: Friday, August 04, 2006 12:05 PM
To: Schulwitz, Paul
Subject: FW: 09/904994 nucleic and amino acid sequence search with specific gap values; claims attached
Importance: High

Paul, After looking at all of the claims, I found that they claims SEQ ID Nos 1, 2 and 3. with the nucleic acid parameters of claim 23 previously emailed, and the polypeptide sequences with a different set of parameters as set forth in emailed claim 34. Thanks!!!

Ginny Portner

Remsen Building

Art Unit 1645

Room E03, B02; Mail Box 3C18

(571) 272-0862

-----Original Message-----

From: Portner, Ginny

Sent: Friday, August 04, 2006 11:44 AM

To: Subject: FW: 09/904994 nucleic and amino acid sequence search with specific gap values; claims attached

Importance: High

Paul, From the Specification, I found that they sued a computer program called "Align Plus for Windows, available from Scientific and Educational Software, P.O. Box 72045, Durham, NC 27722-2045 USA. I hope that this information helps. ginny

Ginny Portner

Remsen Building

Art Unit 1645

Room E03, B02; Mail Box 3C18

(571) 272-0862

-----Original Message-----

From: Portner, Ginny

Sent: Friday, August 04, 2006 11:32 AM

To:Subject: 09/904994 nucleic and amino acid sequence search with specific gap values; claims attached

Importance: High

Thanks for your help with this search issue. I am sending the claims to you now (Snagit training WORKed!!!!!!!) Once I check the Spec for a program name etc I will send you another email. ginny

Ginny Portner

Remsen Building

Art Unit 1645

Room E03, B02; Mail Box 3C18

(571) 272-0862

In the claims:

Claims 1 - 22 (Canceled)

23. (Currently amended) A nucleic acid molecule comprising a nucleotide sequence encoding two ~~urease X and urease Y~~ subunit polypeptides of a urease complex such as expressed by *Helicobacter felis*, said nucleotide sequence having at least 85% homology with SEQ ID NO: 1 or a part thereof encoding at least an immunogenic fragment of one of said subunits, said part having a length of at least 40 ~~70~~ nucleotides.

34. (Currently amended) An isolated *Helicobacter felis* urease X subunit polypeptide, said polypeptide comprising an amino acid sequence that is at least 85% homologous to SEQ ID NO: 2, or an immunogenic fragment of said polypeptide having a length of at least 40 ~~70~~ amino acids, wherein said immunogenic fragment induces an immune response against an urease X subunit in ureaseXY.

Amendments to the Claims:

min length = 70

1. – 22. (cancelled)

23. (currently amended) [[A]] An isolated nucleic acid molecule comprising

(i) a nucleotide sequence encoding urease X and urease Y subunit polypeptides of a urease complex expressed by *Helicobacter felis*, said nucleotide sequence having at least 85% homology with SEQ ID NO: 1 as determined over a global alignment wherein the mismatch value is 1, the open gap value is 3 and the extended gap value is 3; or

(ii) a part ~~thereof~~ of (i) encoding at least an immunogenic fragment of one of said subunits, said part having a length of at least 70 nucleotides.

24. – 25. (cancelled)

26. (original) The nucleic acid molecule of claim 23, which encodes one or both of the urease X subunit polypeptide and urease Y subunit polypeptide.

27. (cancelled)

28. (currently amended) The nucleic acid molecule of claim 23, wherein the nucleotide sequence has at least 94% homology with SEQ ID No: 1 as determined over a global alignment wherein the mismatch value is 1, the open gap value is 3 and the extended gap value is 3.

29. (cancelled)

30. (previously presented) An isolated and purified DNA fragment comprising a nucleotide sequence according to claim 23.

31. (original) A recombinant DNA molecule comprising a nucleotide sequence according to Claim 23 under the control of a functionally linked promoter.

32. (original) A live recombinant carrier comprising a recombinant DNA molecule of claim 31.

33. (original) A host cell comprising a nucleic acid molecule of claim 23, a DNA fragment of claim 30, a recombinant DNA molecule of claim 31 or a live recombinant carrier of claim 32.

34. (currently amended) An isolated *Helicobacter felis* urease X subunit polypeptide, said polypeptide comprising

(a) an amino acid sequence that is at least 85% homologous to SEQ ID NO: 2 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1: [[,]] or

(b) an immunogenic fragment of (a) ~~said polypeptide~~ having a length of at least 70 amino acids, wherein said immunogenic fragment induces an immune response against [[an]] a urease X subunit in ureaseXY.

35. -36. (cancelled)

37. (currently amended) The polypeptide of claim 34, wherein said polypeptide comprises:

(i) an amino acid sequence that is at least 90% homologous to SEQ ID NO: 2 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1: [[,]] or

(ii) an immunogenic fragment of ~~said polypeptide which~~ (i) having a length of at least 70 amino acids;
wherein said immunogenic fragment induces an immune response against [[an]] a urease X subunit in ureaseXY.

38. (currently amended) The polypeptide of claim ~~[[37]]~~ 34, wherein said polypeptide comprises:

(i) an amino acid sequence that is at least 94% homologous to SEQ ID NO: 2 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1; [[,]] or

(ii) an immunogenic fragment of ~~said polypeptide which~~ (i) having a length of at least 70 amino acids;
wherein said immunogenic fragment induces an immune response against [[an]] a urease X subunit in ureaseXY.

39. (currently amended) The polypeptide of claim ~~[[31]]~~ 34, wherein said polypeptide comprises:

(i) an amino acid sequence that is at least ~~[[97%]]~~ 99% homologous to SEQ ID NO: 2 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1; [[,]] or

(ii) an immunogenic fragment of ~~said polypeptide which~~ (i) having a length of at least 70 amino acids;
wherein said immunogenic fragment induces an immune response against [[an]] a urease X subunit in ureaseXY.

40. (currently amended) An isolated *Helicobacter felis* urease Y subunit polypeptide, said polypeptide comprising

(a) an amino acid sequence that is at least ~~[[85%]]~~ 86% homologous to SEQ ID NO: 3 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1; [[,]] or

(b) an immunogenic fragment of (a) ~~said polypeptide~~ having a length of at least 70 amino acids, wherein said immunogenic fragment induces an immune response against [[an]] a ureaseY subunit in ureaseXY.

41.- 43. (cancelled)

44. (currently amended) The polypeptide of claim 40, wherein said polypeptide comprises:

(i) an amino acid sequence that is at least [[94%]] 98% homologous to SEQ ID NO: 3 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1: [[,]] or

(ii) an immunogenic fragment of said polypeptide, which (i) having a length of at least 70 amino acids;

wherein said immunogenic fragment induces an immune response against [[an]] a urease Y subunit in ureaseXY.

45. (cancelled)

46. (previously presented) An immunogenic composition comprising an immunogenically effective amount of a nucleic acid molecule of claim 23 and a pharmaceutically acceptable carrier.

47. (previously presented) The immunogenic composition of claim 46, further comprising an adjuvant.

48. (previously presented) The immunogenic composition of claim 46, further comprising an additional antigen derived from a virus or microorganism which is pathogenic to mammals.

49. (previously presented) The immunogenic composition of claim 48, wherein said virus or microorganism pathogenic to mammals is selected from the group consisting of Feline Infectious Peritonitis virus, Feline Immune deficiency virus, Canine Parvovirus, Feline Parvovirus, Distemper virus, Adenovirus, Calicivirus, *Bordetella bronchiseptica*, *Borrelia burgdorferi*,

Leptospira interrogans, *Chlamydia* and *Bartonella henseli*.

50. (previously presented) An immunogenic composition for combating *Helicobacter felis* infections, comprising antibodies against a polypeptide of claims 34 or 40.

51.- 56. (cancelled)

57. (previously presented) An immunogenic composition, comprising an immunogenically effective amount of a polypeptide according to Claim 34 and a pharmaceutically acceptable carrier.

58. (previously presented) An immunogenic composition, comprising an immunogenically effective amount of a polypeptide according to claim 40 and pharmaceutically acceptable carrier.

59. (new) The polypeptide of claim 40, wherein said polypeptide comprises:

(i) an amino acid sequence that is at least 99% homologous to SEQ ID NO: 3 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1; or

(ii) an immunogenic fragment of (i) having a length of at least 70 amino acids; wherein said immunogenic fragment induces an immune response against a urease Y subunit in ureaseXY.

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2006, 01:13:55 ; Search time 16252 Seconds
(without alignments)
11343.860 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2880.6
Sequence: 1 rragagatttccacactt.....aaaaagtagaacacagg 2883

Scoring table: IDENTITY_NUC
Gapop 3.0 , Gapext 3.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_env.*
2: gb_pat.*
3: gb_ph.*
4: gb_pl.*
5: gb_pr.*
6: gb_ro.*
7: gb_sts.*
8: gb_sy.*
9: gb_un.*
10: gb_vi.*
11: gb_ov.*
12: gb_hvg.*
13: gb_in.*
14: gb_om.*
15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2880.6	100.0	2883	2	BD185302 Helicobac
2	2880.6	100.0	2883	2	AX356683 Sequence
3	2228	77.3	2452	2	BD185306 Helicobac
4	2228	77.3	2452	2	AX356695 Sequence
5	2158.6	74.9	2405	2	BD185303 Helicobac
6	2158.6	74.9	2405	2	AX356686 Sequence
7	2158.6	74.9	2407	2	BD185305 Helicobac
8	2158.6	74.9	2407	2	AX356692 Sequence
9	1948	67.6	2183	2	BD185304 Helicobac
10	1948	67.6	2183	2	AX356689 Sequence
11	1135.6	39.4	8406	15	AF330621 Helicobac
12	1099.8	38.2	2664	15	HECUREASE
13	1040.8	36.1	110000	15	AE001439.00
14	1029.2	35.7	110000	15	AE000511.00
15	1024.4	35.6	5100	15	HECUREABCD
16	1022	35.5	7088	15	AB032429 Helicobac
17	1021.4	35.5	2735	2	AR054309
18	1008.8	35.0	2619	2	A41006 Sequence 1

19	1008.8	35.0	2619	2	A72509	A72509 Sequence 1
20	1008.8	35.0	2619	2	A73285	A73285 Sequence 1
21	1008.8	35.0	2619	2	AR061045	AR061045 Sequence
22	1008.8	35.0	2619	2	AR161678	AR161678 Sequence
23	1008	35.0	2746	15	HPHPUABG	X17079 Helicobacte
24	1008	35.0	2767	2	A31515	A31515 H.pylori ur
25	1005.8	34.9	4824	2	AR349305	AR349305 Sequence
26	1005.8	34.9	4824	2	AR349306	AR349306 Sequence
27	1005.4	34.9	2430	15	AY227442	AY227442 Helicobac
28	1004	34.9	2619	15	HFURE	X69080 H.felis ure
29	963.8	33.5	6126	15	AF332656	AF332656 Helicobac
30	963.8	33.5	110000	15	AR017125_03	Continuation (4 of
31	894.6	31.1	1704	15	AY714224	AY714224 Helicobac
32	885.8	30.8	1815	2	BD092530	BD092530 Identific
33	884.8	30.7	1710	15	AY295085	AY295085 Helicobac
34	881.8	30.6	1707	2	CQ903244	CQ903244 Sequence
35	881.8	30.6	1707	2	CQ903245	CQ903245 Sequence
36	880.2	30.6	1704	15	DO141576	DO141576 Helicobac
37	880	30.5	1710	15	AF352376	AF352376 Helicobac
38	879.8	30.5	1680	15	AF479026	AF479026 Helicobac
39	878.4	30.5	1710	2	A07398	A07398 C.jejuni ge
40	878.4	30.5	1710	2	A08818	A08818 C.jejuni DN
41	776.6	27.0	10118	15	BSUEROP	Y08559 Bacillus su
42	776.6	27.0	200690	15	BSUB0019	Z99122 Bacillus su
43	775.6	26.9	110000	15	BA000004_02	Continuation (3 of
44	714.6	24.8	4723	15	AB201709	AB201709 Campyloba
45	705.2	24.5	110000	15	CP000020_07	Continuation (8 of

ALIGNMENTS

RESULT 1	BD185302	BD185302	2883 bp	DNA	linear	PAT 17-JUN-2003
LOCUS	BD185302	Helicobacter felis vaccine.				
DEFINITION	BD185302	Helicobacter felis vaccine.				
ACCESSION	BD185302.1	GI:31877502				
VERSION	JP 2002355054-A/1.					
KEYWORDS	JP 2002355054-A/1.					
SOURCE	Helicobacter felis					
ORGANISM	Helicobacter felis					
REFERENCE	1 (bases 1 to 2883)					
AUTHORS	Kusters,J.G. and Cattoli,G.					
TITLE	Helicobacter felis vaccine					
JOURNAL	Patent: JP 2002355054-A 1 10-DEC-2002;					
COMMENT	AKZO NOBEL NV					
	OS Helicobacter felis					
	PN JP 2002355054-A/1					
	PD 10-DEC-2002					
	PF 16-JUL-2001 JP 2001214711					
	PR 17-JUL-2000 EP 00202565.8					
	PI JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI					
	PC					
	C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC					
	A61K39/23					
	PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,					
	PC C12N1/19,					
	PC					
	C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC					
	53,					
	PC					
	G01N33/53,G01N33/566,G01N33/569/(C12N9/80,C12R1/01),(C12Q1/68, PC					
	C12R1/01),					
	PC C12N15/00,C12N5/00,A61K37/02					
	CC Helicobacter felis vaccine					
	FF Key Location/Qualifiers					
	FT CDS (206)..(886)					
	FT CDS (897)..(2603).					
	Location/Qualifiers					
	1..2883					
	/organism="Helicobacter felis"					

FEATURES
source

/mol type="genomic DNA" /db_xref="taxon:214"									
ORIGIN									
Query Match 100.0%; Score 2880.6; DB 2; Length 2883;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	RGGRAGATTTCCARCACTTCCAGCACATATTGATCCTGTGTGGTGGTGAATTCR	60						
Db	1	RGGRAGATTTTCARCACTTCCAGCACATATTGATCCTGTGTGGTGGTGAATTCR	60						
Qy	61	ACTTGTAAATCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA	120						
Db	61	ACTTGTAAATCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA	120						
Qy	121	TTACTTATTAATAAAGTTAATAAAGTAACGAAATTAAGACTATAATCCATTGCCTT	180						
Db	121	TTACTTATTAATAAAGTTAATAAAGTAACGAAATTAAGACTATAATCCATTGCCTT	180						
Qy	181	TAAATTTAAACAAGAGTAAATAGTGAATCATCAACCCAAAGAGCAAGAAAGTCTT	240						
Db	181	TAAATTTAAACAAGAGTAAATAGTGAATCATCAACCCAAAGAGCAAGAAAGTCTT	240						
Qy	241	GTTATATTATCGGCGGAAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACA	300						
Db	241	GTTATATTATCGGCGGAAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACA	300						
Qy	301	ACCGGAAGCCATTCCTACATTTAGTGCCCATATTATGACGAAGCGCGCTGGAAGAAA	360						
Db	301	ACCGGAAGCCATTCCTACATTTAGTGCCCATATTATGACGAAGCGCGCTGGAAGAAA	360						
Qy	361	AACCGTTGCCAGCTTATGGAAGAGTGCATGCACATTTTGAAGAAAGATGAATGCC	420						
Db	361	AACCGTTGCCAGCTTATGGAAGAGTGCATGCACATTTTGAAGAAAGATGAATGCC	420						
Qy	421	CGGGGTGGTAAATATGTTCCCGATCTAGGTGTAGAGCCACCTTCTGATGTACGAA	480						
Db	421	CGGGGTGGTAAATATGTTCCCGATCTAGGTGTAGAGCCACCTTCTGATGTACGAA	480						
Qy	481	ACTTGTAACTGTGAATTTGGCCCATCGAACACAGATGAGCACCTTCAAAGCGGCGAAGTCAA	540						
Db	481	ACTTGTAACTGTGAATTTGGCCCATCGAACACAGATGAGCACCTTCAAAGCGGCGAAGTCAA	540						
Qy	541	ATTTGGTTGCGATAAAGACATCGAGCTCAATCGAGGCAAGAAAGTAAACCGAACTTGAGGT	600						
Db	541	ATTTGGTTGCGATAAAGACATCGAGCTCAATCGAGGCAAGAAAGTAAACCGAACTTGAGGT	600						
Qy	601	TACTAAAGAAGGCTAAATCCTTGATGTGGGTAGCCATTTCCACTTTTGAAGCTAA	660						
Db	601	TACTAAAGAAGGCTAAATCCTTGATGTGGGTAGCCATTTCCACTTTTGAAGCTAA	660						
Qy	661	CAAGGCACTAAATTCGATCTGAAAGAGCTATGGCAAGCGCTAGATATTCCTCTGG	720						
Db	661	CAAGGCACTAAATTCGATCTGAAAGAGCTATGGCAAGCGCTAGATATTCCTCTGG	720						
Qy	721	CAACACGCTACGATTTGGGGCAGGACAAACCGCAAGTGCAGTTGATTCCTCTGGTGG	780						
Db	721	CAACACGCTACGATTTGGGGCAGGACAAACCGCAAGTGCAGTTGATTCCTCTGGTGG	780						
Qy	781	CAGTAAATAAGTATGTCATGAACCGGCTTGTGAATAACATCGCGGATGAACGCCATAA	840						
Db	781	CAGTAAATAAGTATGTCATGAACCGGCTTGTGAATAACATCGCGGATGAACGCCATAA	840						
Qy	841	ACATAAGCGCTTCACAGGCGAAATCTCAGGATTTATCAAGTAAGGAGACTCCCATGA	900						
Db	841	ACATAAGCGCTTCACAGGCGAAATCTCAGGATTTATCAAGTAAGGAGACTCCCATGA	900						
Qy	901	AAATGAAATAAACAAGAAATATGTAATACCTACGGAACCCCAAGAGCGGATGAAGTGCCT	960						
Db	901	AAATGAAATAAACAAGAAATATGTAATACCTACGGAACCCCAAGAGCGGATGAAGTGCCT	960						
Qy	961	TAGGAGATACCGATCTTTGGGCAGAGTAGAACAATGACTATACCACTTATGGCGAAGAAC	1020						

Db	961	TAGGAGATACCGATCTTTGGGCAGAGTAGAACAATGACTATACCACTTATGGCGAAGAAC	1020						
Qy	1021	TTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGTGAGGCAATAGCCCTGATG	1080						
Db	1021	TTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGTGAGGCAATAGCCCTGATG	1080						
Qy	1081	AAAAACACCTAGATTTAGTCATCACTAACCGCATGATTAATCGATACACGGGATTTACA	1140						
Db	1081	AAAAACACCTAGATTTAGTCATCACTAACCGCATGATTAATCGATACACGGGATTTACA	1140						
Qy	1141	AAGCCGACATTTGGGATTAATAACGGCAAAATCCATGGCATTTGGCAAGGAGGAAACAAG	1200						
Db	1141	AAGCCGACATTTGGGATTAATAACGGCAAAATCCATGGCATTTGGCAAGGAGGAAACAAG	1200						
Qy	1201	ACATGCAAGATGGCGCTTAAGCCCTCATATGGTCTGGGTGTGGGCACAGAAGCACTAGCAG	1260						
Db	1201	ACATGCAAGATGGCGCTTAAGCCCTCATATGGTCTGGGTGTGGGCACAGAAGCACTAGCAG	1260						
Qy	1261	GGGAAGGTATGATTAATTAACCGCTGGGGGAATCGATTCACACACCACTTCTTTCTCCAC	1320						
Db	1261	GGGAAGGTATGATTAATTAACCGCTGGGGGAATCGATTCACACACCACTTCTTTCTCCAC	1320						
Qy	1321	AACAAATCCCTACCGCTCTAGCCCAATGGCTTACAAACCATTTTGGAGGCGGCACAGTTC	1380						
Db	1321	AACAAATCCCTACCGCTCTAGCCCAATGGCTTACAAACCATTTTGGAGGCGGCACAGTTC	1380						
Qy	1381	CTGTAGATGGCAAGCAATCGACTATCTACCTCCGGGCAATGGAACTTGCACCCGATGT	1440						
Db	1381	CTGTAGATGGCAAGCAATCGACTATCTACCTCCGGGCAATGGAACTTGCACCCGATGT	1440						
Qy	1441	TGCGCGCAGCAGAGAGATTTCTTGAATGTGGGCTTTTGGGCAAGCAATAGCTCTA	1500						
Db	1441	TGCGCGCAGCAGAGAGATTTCTTGAATGTGGGCTTTTGGGCAAGCAATAGCTCTA	1500						
Qy	1501	GCMAAAACAACTTGTAGAACAGTAGAAGCGGGCGCATTTGGTTTAAATTCATGAAG	1560						
Db	1501	GCMAAAACAACTTGTAGAACAGTAGAAGCGGGCGCATTTGGTTTAAATTCATGAAG	1560						
Qy	1561	ACTGGGCGACAAACCAAGTCGATCTGCTTGTAGCGTGGCAGATGAATACGATG	1620						
Db	1561	ACTGGGCGACAAACCAAGTCGATCTGCTTGTAGCGTGGCAGATGAATACGATG	1620						
Qy	1621	TGCAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGTTTATGTAGATGACACCTTAA	1680						
Db	1621	TGCAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGTTTATGTAGATGACACCTTAA	1680						
Qy	1681	ATGCATGAACGGGCGCCATCCATGCTTACCACTTGGAGGAGCGGTGGAGGACACT	1740						
Db	1681	ATGCATGAACGGGCGCCATCCATGCTTACCACTTGGAGGAGCGGTGGAGGACACT	1740						
Qy	1741	CACCTGATTTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCCA	1800						
Db	1741	CACCTGATTTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCCA	1800						
Qy	1801	CTATTCCCTATACCAATTAATACGGTTGCAGAACACTTAGACATGCTCATGATGCCACC	1860						
Db	1801	CTATTCCCTATACCAATTAATACGGTTGCAGAACACTTAGACATGCTCATGATGCCACC	1860						
Qy	1861	ACCTAGACAAACGCAATCCGAGGAGATTTACAAATTTTCTCAAGCGGTATCCGCCCGGCT	1920						
Db	1861	ACCTAGACAAACGCAATCCGAGGAGATTTACAAATTTTCTCAAGCGGTATCCGCCCGGCT	1920						
Qy	1921	CTATCGCGCTGAGATGTCTCCATGATATGGGTGTGATCCGATGACAAGCTCGGATTT	1980						
Db	1921	CTATCGCGCTGAGATGTCTCCATGATATGGGTGTGATCCGATGACAAGCTCGGATTT	1980						
Qy	1981	CGCAAGCAATGGGCGCTGAGGCGGAGTGTCTCGAACTTTGGCAGACTGCGGATAAGA	2040						
Db	1981	CGCAAGCAATGGGCGCTGAGGCGGAGTGTCTCGAACTTTGGCAGACTGCGGATAAGA	2040						
Qy	2041	ATAAAAAGAAATTTGGTGAAGCTTCTGAAGATGGCAAGATGAACGATAATTTCCGCAITTA	2100						

Db	2041	ATAAAAAGAAATTGGTAAGCTTCTGGAAGTGCAGGATGCAAGATACGATAATTTCCGCATT	2100
Qy	2101	AGCGTACATCTCTCAAAATACATATCAACCCCGCTTTGACCCACGGCGTGAGCGAGTATA	2160
Db	2101	AGCGTACATCTCTCAAAATACATATCAACCCCGCTTTGACCCACGGCGTGAGCGAGTATA	2160
Qy	2161	TCGGCTCTGTGGAGAGGCAAGATCCGCACTTGTGTGGTGTGGAATCCTGCTTTTGTG	2220
Db	2161	TCGGCTCTGTGGAGAGGCAAGATCCGCACTTGTGTGGTGTGGAATCCTGCTTTTGTG	2220
Qy	2221	GCCTAAACCCCAAAATCGTGATCAAGCGGTATGTGTCTTCTCTGAAATGGCGATT	2280
Db	2221	GCCTAAACCCCAAAATCGTGATCAAGCGGTATGTGTCTTCTCTGAAATGGCGATT	2280
Qy	2281	CTAACCGCTCTGTGCCCATCTCCCAACCGGTTTATACCGCAAAATGTTGGGCATCAG	2340
Db	2281	CTAACCGCTCTGTGCCCATCTCCCAACCGGTTTATACCGCAAAATGTTGGGCATCAG	2340
Qy	2341	GCAAGCGAAATTGTACACAGCATCACTTTGTTTCCAAAGTCGCCCTATGAAATGGCG	2400
Db	2341	GCAAGCGAAATTGTACACAGCATCACTTTGTTTCCAAAGTCGCCCTATGAAATGGCG	2400
Qy	2401	TGAAGAAAGCTGGGCTTAGAGCGCAAGTCTTACCGGTCAAAACTGCGGTAAACATCA	2460
Db	2401	TGAAGAAAGCTGGGCTTAGAGCGCAAGTCTTACCGGTCAAAACTGCGGTAAACATCA	2460
Qy	2461	CAAGAAAGACTTCAAGTTCAACGACAAACCGCAAAATCACCGTCGATCCGAAACCT	2520
Db	2461	CAAGAAAGACTTCAAGTTCAACGACAAACCGCAAAATCACCGTCGATCCGAAACCT	2520
Qy	2521	TCAGGTCTTTGTAGATGGCAACTCTGCACTTAAACCCACCTCGCAAGTGCTCTAG	2580
Db	2521	TCAGGTCTTTGTAGATGGCAACTCTGCACTTAAACCCACCTCGCAAGTGCTCTAG	2580
Qy	2581	CCAGCGCTACCTTCTTAGGCAAAATGCCCCCTTTGGGGCAGGTTATTAGGAA	2640
Db	2581	CCAGCGCTACCTTCTTAGGCAAAATGCCCCCTTTGGGGCAGGTTATTAGGAA	2640
Qy	2641	TCCTCATCAACCGCACCTGCAATCGGCTTTCGCTGCGGATCGTGCTTAAACCAAC	2700
Db	2641	TCCTCATCAACCGCACCTGCAATCGGCTTTCGCTGCGGATCGTGCTTAAACCAAC	2700
Qy	2701	TTTTCATCTTTAAGCAATGCCCAATTTTAAATTAATTAATCTTATTAATTAATATAT	2760
Db	2701	TTTTCATCTTTAAGCAATGCCCAATTTTAAATTAATTAATCTTATTAATTAATATAT	2760
Qy	2761	TATGCCCTCATTTTAAAGGAGAAATATGCGTAGTCTTTGCTATGCTATGCGGGG	2820
Db	2761	TATGCCCTCATTTTAAAGGAGAAATATGCGTAGTCTTTGCTATGCTATGCGGGG	2820
Qy	2821	TTGTTTGGTCTGGCGCAAGGATTTGAAACCCATCGCTCAAAAAAGTAGAAGCCAC	2880
Db	2821	TTGTTTGGTCTGGCGCAAGGATTTGAAACCCATCGCTCAAAAAAGTAGAAGCCAC	2880
Qy	2881	AGG 2883	
Db	2881	AGG 2883	

RESULT 2
AX356683
LOCUS AX356683
DEFINITION Sequence 1 from Patent EP1176192.
ACCESSION AX356683
VERSION AX356683.1 GI:18674020
KEYWORDS Helicobacter felis
SOURCE Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
ORGANISM Helicobacteraceae; Helicobacter.
REFERENCE 1
AUTHORS Kusters, J.G. and Cattoli, G.
TITLE Helicobacter felis vaccine

JOURNAL	Patent: EP 1176192-A 1 30-JAN-2002;
FEATURES	Akzo Nobel N.V. (NL)
source	Location/Qualifiers
	1. 2883
	/organism="Helicobacter felis"
	/mol_type="unassigned DNA"
	/db_xref="taxon:214"
CDS	206..886
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD23255.1"
	/db_xref="GI:18674021"
	/translation="MKLTPKQEKFLLYAGEVARKRKABGLKLNQPEALAIYSAIHM
	DEARRGKKTVAQLMECHFLKQDVMFGVMGPDGLGVEATFPDGTKLVTVMVLEP
	DEHFKAGEVFKGCDKIDELNAGKEVTEVNTGPKSLHVGSHFHFPEANKLKFRE
	KAYGKRLDIPSGNTLRIGAGQTRKVLIPLGSSKKVIGMNLVNNIADERHKHAKLDK
	AKSHGFIK"
CDS	897..2603
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD23256.1"
	/db_xref="GI:18674022"
	/translation="MKMKKQBYNTYGTGDKVRLGDTDLWAEVHDYTTVEELKF
	GAGKTIREGMGQNSPDENLIDLVTINAMIIDYTGIIYKADI GIKNGKIHGIGAGNKG
	MDQGVSMVVGVEALAGGEMITAGGIDSHFLSPQFPPTALANGVITMFGGKT
	PGVDGTNATITPGKWNLRMLRAEYSMNVFLGKGNSSKQIOLVEVGAIGFGK
	LHEDWGTTPSAIDHCLSVADRYDVQVCIHTDTVNEAGYVDDTLNMGNGRAIHAYHLEG
	AGGSHSDVITMAGELNLPSTTPTIPIYNTINTVAEHLDMLTCHILDKRILREDLQFS
	QSRIRPSGIAADVLHDMGVJAMTSSDSQAMGRAGEVIRPTWQTDADNKKKEFKLPED
	KQNDNFRIKRYISKYTINPALTHGVSYIGSVEEGKADLVNMPAFVGVKPKIVIK
	GMVVFSEMGDSNASVPTQPVYREMFHGKAKFDTISITFVSKVAYENGVEKELGL
	ERQVLPVKNCRNITKDPFKNDKTAKITVDPKTFEVDGKLC TSKPTSOVPLAQRYT
	FF"

ORIGIN

Query Match	100.0%; Score 2880.6; DB 2; Length 2883;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2883; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 RGRAGATTTTCCARCACTTCAAGCACAATATGATCCTGTGTGTGGGTGGAATTCR 60
Db	1 RGRAGATTTTCCARCACTTCAAGCACAATATGATCCTGTGTGTGGGTGGAATTCR 60
Qy	61 ACTTGTAACTCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA 120
Db	61 ACTTGTAACTCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA 120
Qy	121 TTACTTATATTAATAAGTTAATAAAGTAACGAAATTAGGACTATAATCCCATTGCCTT 180
Db	121 TTACTTATATTAATAAGTTAATAAAGTAACGAAATTAGGACTATAATCCCATTGCCTT 180
Qy	181 TAAATTTAAACAAGGAGTAATAGTGAACCTCACACCAAGAGCAAGAAAGTTCTT 240
Db	181 TAAATTTAAACAAGGAGTAATAGTGAACCTCACACCAAGAGCAAGAAAGTTCTT 240
Qy	241 GTTATATATGCGGCGAAGTGGCTAGAAAGCGCAAGACGAGGCGCTTAAAGCTCAACCA 300
Db	241 GTTATATATGCGGCGAAGTGGCTAGAAAGCGCAAGACGAGGCGCTTAAAGCTCAACCA 300
Qy	301 ACCCGAAGCATTCTTACATTTAGTGGCCCATATTATGACGAAGCGCGCGCGGGAAGAAA 360
Db	301 ACCCGAAGCATTCTTACATTTAGTGGCCCATATTATGACGAAGCGCGCGCGGGAAGAAA 360
Qy	361 AACCGTTGCCAGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCC 420
Db	361 AACCGTTGCCAGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCC 420
Qy	421 CGGGGTGGGTAAATATGTTTCCCGATCTAGGTGTAGAACCCACCTTTTCTGTATGGTACGAA 480
Db	421 CGGGGTGGGTAAATATGTTTCCCGATCTAGGTGTAGAACCCACCTTTTCTGTATGGTACGAA 480

QY 481 ACTTGTAACCTGTGAATTGGCCCATCGAACCGATGAGCACTTCAAAGCGGGCGAAGTGAA 540
DB 481 ACTTTGTAACCTGTGAATTGGCCCATCGAACCGATGAGCACTTCAAAGCGGGCGAAGTGAA 540
QY 541 ATTTGGTTGGCATAAAGACATCGAGCTCAATCGCAGGCAAGAGTAACCGCAACTTGAGGT 600
DB 541 ATTTGGTTGGCATAAAGACATCGAGCTCAATCGCAGGCAAGAGTAACCGCAACTTGAGGT 600
QY 601 TACTAAATGAAGGGCCCTAAATCTTGCATGTGGGTAGCCATTTCACACTTCTTTGAAGCTAA 660
DB 601 TACTAAATGAAGGGCCCTAAATCTTGCATGTGGGTAGCCATTTCACACTTCTTTGAAGCTAA 660
QY 661 CAAGGCACTAAAATTCGATCGTGAAGGCTATGCGCAAGCCCTAGATATTCCTCTGG 720
DB 661 CAAGGCACTAAAATTCGATCGTGAAGGCTATGCGCAAGCCCTAGATATTCCTCTGG 720
QY 721 CAACAGCTACGCAATGGGGGAGGAGCAACCGCAAGTGCAGTTGATTCCTCTGGTGG 780
DB 721 CAACAGCTACGCAATGGGGGAGGAGCAACCGCAAGTGCAGTTGATTCCTCTGGTGG 780
QY 781 CAGTAAAAAAGTGAATGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAA 840
DB 781 CAGTAAAAAAGTGAATGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAA 840
QY 841 ACATAAAGCGCTTGAACAGGGCAAAATCTCACGGATTTATCAAGTAAGAGACTCCCATGA 900
DB 841 ACATAAAGCGCTTGAACAGGGCAAAATCTCACGGATTTATCAAGTAAGAGACTCCCATGA 900
QY 901 AAATGMAAACAAGAAATGTAAATACCTACGGACCCACCAAGGCGATAAAGTGGCT 960
DB 901 AAATGMAAACAAGAAATGTAAATACCTACGGACCCACCAAGGCGATAAAGTGGCT 960
QY 961 TAGGAGATACCGATCTTTGGGCGAAGTAGAACATGACTATACCACTATGGCGAAGAAC 1020
DB 961 TAGGAGATACCGATCTTTGGGCGAAGTAGAACATGACTATACCACTATGGCGAAGAAC 1020
QY 1021 TTAATTTGGCGCGGTTAAAACTATCCGTGAGGTTATGGGTACAGAGCAATAGCCCTGATG 1080
DB 1021 TTAATTTGGCGCGGTTAAAACTATCCGTGAGGTTATGGGTACAGAGCAATAGCCCTGATG 1080
QY 1081 AAAACACCTTAGATTAGTCACTAATACCGGATGATTTATCGACTACACCGGATTTACA 1140
DB 1081 AAAACACCTTAGATTAGTCACTAATACCGGATGATTTATCGACTACACCGGATTTACA 1140
QY 1141 AAGCCGATTTGGGATTTAAAAACGCAAAATCCATGGCAATGGCAAGCAGGAAACAAG 1200
DB 1141 AAGCCGATTTGGGATTTAAAAACGCAAAATCCATGGCAATGGCAAGCAGGAAACAAG 1200
QY 1201 ACATGCAAGATGGCGTAAGCCCTCATATGGTGTGGGTATGGGCAAGAGCACTAGCAG 1260
DB 1201 ACATGCAAGATGGCGTAAGCCCTCATATGGTGTGGGTATGGGCAAGAGCACTAGCAG 1260
QY 1261 GGGAAAGTATGATTTAATACCGTGGGGAAATCGATTCAACCCACTTCTTTCTCCAC 1320
DB 1261 GGGAAAGTATGATTTAATACCGTGGGGAAATCGATTCAACCCACTTCTTTCTCCAC 1320
QY 1321 AACAAATCCCTACCGCTCTAGCCATGCGTTTACAAACCATGTTTGGAGCGCGCACAGTTC 1380
DB 1321 AACAAATCCCTACCGCTCTAGCCATGCGTTTACAAACCATGTTTGGAGCGCGCACAGTTC 1380
QY 1381 CTGTAGATGGCAAGAAATCGCACTATCATCTCCGGGCAAAATGGAACTTGGACCCGATGT 1440
DB 1381 CTGTAGATGGCAAGAAATCGCACTATCATCTCCGGGCAAAATGGAACTTGGACCCGATGT 1440
QY 1441 TGGCGGACGAGCAAGAGATTTCTATGAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTA 1500
DB 1441 TGGCGGACGAGCAAGAGATTTCTATGAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTA 1500
QY 1501 GCMAAACAACCTTGTAGAACAGTAGAGCGGCGGATTTGTTTAAATTTGCATGAG 1560
DB 1501 GCMAAACAACCTTGTAGAACAGTAGAGCGGCGGATTTGTTTAAATTTGCATGAG 1560
QY 1561 ACTGGGGGCAACACCAAGTGGATCGATCACTGTGTGAGCGTGGCAGATGAATACGATG 1620

DB 1561 ACTGGGGGCAACACCAAGTGGATCGATCACTGTGTGAGCGTGGCAGATGAATACGATG 1620
QY 1621 TGCAGTTTGTATCCACACCGGATCAGTCAATGAGGACAGTTATGTAGATGACACCCCTAA 1680
DB 1621 TGCAGTTTGTATCCACACCGGATCAGTCAATGAGGACAGTTATGTAGATGACACCCCTAA 1680
QY 1681 ATGCAATGAACGGGCGCGCATCCATCGCTACCAATTTGAGGGAGCGGTGGAGGACACT 1740
DB 1681 ATGCAATGAACGGGCGCGCATCCATCGCTACCAATTTGAGGGAGCGGTGGAGGACACT 1740
QY 1741 CACCTGATTTATCACCATGCGAGCGAGCTCAATTTCTACCTCTCCACACCCCA 1800
DB 1741 CACCTGATTTATCACCATGCGAGCGAGCTCAATTTCTACCTCTCCACACCCCA 1800
QY 1801 CTATTCCCTATACCAATTAATACCGTTGCAAGACATTTAGACATGCTCATGACATGCAACC 1860
DB 1801 CTATTCCCTATACCAATTAATACCGTTGCAAGACATTTAGACATGCTCATGACATGCAACC 1860
QY 1861 ACCTAGACAAAGCGCATCCGCGAGGATTTACAATTTTCTCAAGCCGTATCCGCCCGGCT 1920
DB 1861 ACCTAGACAAAGCGCATCCGCGAGGATTTACAATTTTCTCAAGCCGTATCCGCCCGGCT 1920
QY 1921 CTATCGCGCTGGAAGATGTCTCATGATATGGGTGTGATCGCGATGCAAGCTCGGAT 1980
DB 1921 CTATCGCGCTGGAAGATGTCTCATGATATGGGTGTGATCGCGATGCAAGCTCGGAT 1980
QY 1981 CGCAAGCAATGGGGCGTGCGAGCGAAGTATTCCTCGAACTTTGGGACACTTGGCGATAAGA 2040
DB 1981 CGCAAGCAATGGGGCGTGCGAGCGAAGTATTCCTCGAACTTTGGGACACTTGGCGATAAGA 2040
QY 2041 ATAAAAAGAAATTTGGTAAGCTTCCCTGAAGATGGCAAGATAACGATAATTTCCGCAATTA 2100
DB 2041 ATAAAAAGAAATTTGGTAAGCTTCCCTGAAGATGGCAAGATAACGATAATTTCCGCAATTA 2100
QY 2101 AGCGCTCATCTCCTCAAAATACATCAACCCCGCTTTGACCCAGCGCGTGGAGGATTA 2160
DB 2101 AGCGCTCATCTCCTCAAAATACATCAACCCCGCTTTGACCCAGCGCGTGGAGGATTA 2160
QY 2161 TCGGCTCTGTGGAAGAGGCAAGATCGCGGACTTGGTGTGGAAATCTGCTCTTTTTC 2220
DB 2161 TCGGCTCTGTGGAAGAGGCAAGATCGCGGACTTGGTGTGGAAATCTGCTCTTTTTC 2220
QY 2221 GCGTAAAAACCCAAATTCGTGATCAAGGCGGTATGGTGTCTTCTCTGAAATGGGCGATT 2280
DB 2221 GCGTAAAAACCCAAATTCGTGATCAAGGCGGTATGGTGTCTTCTCTGAAATGGGCGATT 2280
QY 2281 CTAACGCGTCTGTGCGCACTCCCAACCGGTTTATACCGCGAAATGTTTGGGCATCACG 2340
DB 2281 CTAACGCGTCTGTGCGCACTCCCAACCGGTTTATACCGCGAAATGTTTGGGCATCACG 2340
QY 2341 GCAAGCGGAAATTTGACACCGAGCATCACTTTGTTTCCAAAGTCGCTATGAAAATGGCG 2400
DB 2341 GCAAGCGGAAATTTGACACCGAGCATCACTTTGTTTCCAAAGTCGCTATGAAAATGGCG 2400
QY 2401 TGAAGAAAAAGCTGGGCTTAGAGCGCAAGTTCACCGGTCAAAAACTGCGCTAAACATCA 2460
DB 2401 TGAAGAAAAAGCTGGGCTTAGAGCGCAAGTTCACCGGTCAAAAACTGCGCTAAACATCA 2460
QY 2461 CCAAGAAAGACTTCAAGTTTCAACGACAAACCGCAAAATCACCGTCGATCCGAAACCT 2520
DB 2461 CCAAGAAAGACTTCAAGTTTCAACGACAAACCGCAAAATCACCGTCGATCCGAAACCT 2520
QY 2521 TCGAGCTCTTGTAGATGGCAAACTCTGCACTCTTAAACCCACCTCGCAAGTGCCTCTAG 2580
DB 2521 TCGAGCTCTTGTGTAGATGGCAAACTCTGCACTCTTAAACCCACCTCGCAAGTGCCTCTAG 2580
QY 2581 CCCAGCGCTACACTTTCTTCTAGGCAATGCCCCCTTTGGGGGCGAGGTTATTTTAGGAA 2640
DB 2581 CCCAGCGCTACACTTTCTTCTAGGCAATGCCCCCTTTGGGGGCGAGGTTATTTTAGGAA 2640
QY 2641 TCTTCATCAACGCACTTGTGCAATCGGCTTGTGGGTGTGCGATCGTGTCTTTAAAAACAAC 2700

```

Db 2641 TCTTCATCAACAGCACCTGCAATCGTCTTGGTGTGCGATCGTGTGCTTTTAAACACAC 2700
Qy 2701 TTTTCATCTTTAAGCAATCGCCATTTTAAATTAATTAATCTTAAATTAATATATATAT 2760
Db 2701 TTTTCATCTTTAAGCAATCGCCATTTTAAATTAATTAATCTTAAATTAATATATATAT 2760
Qy 2761 TATGCCCTCATTTTAAAGGAGAAATATGCTAGTCTTTGGTATTTGCTATGTGGGG 2820
Db 2761 TATGCCCTCATTTTAAAGGAGAAATATGCTAGTCTTTGGTATTTGCTATGTGGGG 2820
Qy 2821 TTGTTTGGTGTGCGCGCAAGAGGTATTTGAAACCCATCGCCTCAAAAAGTAGAAGCCAC 2880
Db 2821 TTGTTTGGTGTGCGCGCAAGAGGTATTTGAAACCCATCGCCTCAAAAAGTAGAAGCCAC 2880
Qy 2881 AGG 2883
Db 2881 AGG 2883

RESULT 3
BD185306
LOCUS BD185306 2452 bp DNA linear PAT 17-JUN-2003
DEFINITION Helicobacter felis vaccine.
ACCESSION BD185306
VERSION BD185306.1 GI:31877506
KEYWORDS JP 2002355054-A/5.
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 2452)
AUTHORS Kusters,J.G. and Cattolli,G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: JP 2002355054-A 5 10-DEC-2002;
AKZO NOBEL NV
COMMENT OS Helicobacter felis
PN JP 2002355054-A/5
PD 10-DEC-2002
PF 16-JUL-2001 JP 2001214711
PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLLI
PC C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC
A61K39/23,
PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,
PC C12N1/19,
PC C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC
53,
PC G01N33/53,G01N33/566,G01N33/569/(C12N9/80,C12N1/01),(C12Q1/68, PC
C12N1/01),
PC C12N15/00,C12N5/00,A61K37/02
CC Helicobacter felis vaccine
FH Key Location/Qualifiers
FT CDS (48)..(728)
FT CDS (739)..(2445).
FEATURES
source 1..2452
/organism="Helicobacter felis"
/mol_type="genomic DNA"
/db_xref="taxon:214"

ORIGIN
Query Match 77.3%; Score 2228; DB 2; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 159 AGGATATATCCCATGCTTAAATTTAAACAAAGGAGTAATAGTGAATCTACAC 218
Db 1 AGGATATATCCCATGCTTAAATTTAAACAAAGGAGTAATAGTGAATCTACAC 60
Qy 219 CCAAAGAGCAAGAAAGTTCTTGTTATATTATGCGGCGAAGTGGCTTAGAAGCGCAAAG 278

```

```

Db 61 CCAAGAGCAAGAAAGTTCTTGTATATATGCGGCGAAGTGGCTTAGAAGCGCAAAG 120
Qy 279 CAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGTGCCTATTTATGG 338
Db 121 CAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGTGCCTATTTATGG 180
Qy 339 ACGAAGCGCGCGTGGAAAAAAAACCGTTGGCCAGCTTATGGAAGAGTGCATGCACATTTT 398
Db 181 ACGAGCGCGTGTGCGCAAAAAAACCGTTGGCGAACTTATGGAAGAGTGTATGCACATTTT 240
Qy 399 TCAGAAAGATCAAGTAATGCGCGGTGGTAAATATGTTCCGATCTTAGTGTAGAAG 458
Db 241 TCAAAAAAGACGAGGTGATGCGCGGTGGGAATATGTTCCCTGATTTGGGCGTGAAG 300
Qy 459 CCACCTTTCTGATGTCAGAAACTTGTAACTGTGAAATTTGGCCCATCGAACACAGATGAGC 518
Db 301 CCACCTTTCCCGATGGCACCAAACTCTGTAATCTGTGAATTTGGCCCATCGAACATGATGAC 360
Qy 519 ACTTCAAGCGGCGAAGTGAATTTGTTGCGATAAAGACATCGAGCTCAATTCGAGGCA 578
Db 361 ACTTTAAGCGGTGAAGTGAATTTGGCTGTGATTAAGACATTTGAACCTCAACGCGAGTA 420
Qy 579 AAGAAGTAACCGAACTTGAGGTACTATGAAGGGCTTAAATCTTGCATGTGGGTAGCC 638
Db 421 AGGAAGTTTACCGAACTTAGAAGTTACTAACGAAGGACCTTAAATCTTTCATGTGGGTAGCC 480
Qy 639 ATTTCCACTTTTCAAGCTAACAGGCTAAAGCTAAATTTGATCGTGAAGAAAGCCCTATGGCA 698
Db 481 ATTTCCACTTTTGAAGCCAAAGCAATGAAATTCGATCGGAAAGAAAGCCCTATGGCA 540
Qy 699 AACGCTAGATATTCCTCTGCGCAACACGCTACGCAATTTGGGCGAGGACAAACCGCAAAG 758
Db 541 AACGCTAGATATTCCTCTGCGCAACACACTAGCGCAATTTGGGCGAGGACAAACCGTAAAG 600
Qy 759 TGCAGTTGATTCCTTGTGGTGCAGTAAAGAGTGAATTTGGCATGAACGGGCTTTGTAATA 818
Db 601 TGCAGTTAATCCCTCTTGTGGGTAGTAAAAAGTGAATTTGGCATGAACGGGCTTTGTAATA 660
Qy 819 ACATCGCGGATGAAGCCATAAACATAAAGCGCTTGCACAAAGCGGAAATCTCAGCGATTTA 878
Db 661 ATATTGCGGAGAGCGCCATAAACATAAAGCGCTTAGACAAAGCGCTAGACAAAGAAATCTCAGCGATTTA 720
Qy 879 TCAAGTAAGGAGACTCCCATGAAATGAAAAACAAGAAATATGTAATATCTACGAGCAACC 938
Db 721 TCAGTAAGGAGACTCCCATGAAATGAAAAACAAGAGTATGTAATATCTACGAGCAACC 780
Qy 939 ACCAAAGCGATAAAGTGCCTTAGGAGATACCGATCTTTTGGGCGAGAGTAGAACAATGAC 998
Db 781 ACCACAGCGGATAAAGTGCCTTAGGAGATACCGATCTTTTGGGCGAGAGTAGAACAATGAC 840
Qy 999 TATACCACCTATGGCGAAGAACTTAAATTTGGCGGGGTAAAACTATCCGTGAGGGTATG 1058
Db 841 TATACCACCTATGGCGAAGAACTCAAAATTCGCTGAGGTAAAACTATCCGTGAGGGTATG 900
Qy 1059 GGTGAGCAATAGCCCTGATGAAACACCCCTAGATTTAGTATCTACATAACGCGATGATT 1118
Db 901 GGTGAGCAATAGCCCGAGATGAAACACCTTAGATTTAGTATCTACCAACGCGATGATT 960
Qy 1119 ATCGACTACACCGGGATTTACAAAGCGGACATTTGGGATTTAAAAACGCGCAAAATTCATGGC 1178
Db 961 ATTGACTACACCGGGATTTACAAAGCGGACATTTGGCATTTAAAAATGGCAAAATTCATGGC 1020
Qy 1179 ATTGGCAGGCGAGGAAACAAGACATGCAAGATGGCGGTAAAGCCCTCATATGTCGTGGGT 1238
Db 1021 ATTGGCAGGCGAGGAAACAAGACATGCAAGATGGCGGTAAAGCCCTCATATGTCGTGGGT 1080
Qy 1239 GTGGGCGACAGAGCACTAGCAGGCGAAGTATGATTATTACCGCTGGGGGAATTCGATTCA 1298
Db 1081 GTGGGCGACAGAGCACTAGCAGGCGAAGTATGATTATTACCGCTGGGGGAATTCGATTCA 1140
Qy 1299 CACACCCACTTCTCTTCTTCTCCACAAATTCCTACCGCTCTAGCCAAATGGCGCTTACAAC 1358

```

Db 1141 CACACCCACTTCTCTCCACAACAATTCCTACCGCTCTAGCCAATGGCGTTACAACA 1200
Qy 1359 ATGTTTGAGGCGGCACAGGCTCTGTAGATGCGACGAATGCGACTTATCATCTCCGGGC 1418
Db 1201 ATGTTTGCGGTGGCACAGGCCCGGTAGATGCGACGAATGCGACTTACCATCACTCCGGGC 1260
Qy 1419 AAATGGAACTTGCACCGCATGTTGCGCGCAGCAGAGAGTATTCTATGAATGTGGGCTTT 1478
Db 1261 AAATGGAACTTGCACCGCATGTTGCGCGCAGCAGAGAGTATTCTATGAATGTGGGCTTT 1320
Qy 1479 TTGGCAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAACAGTAGTAGAGCGGCGG 1538
Db 1321 TTGGCAAGGCAATAGCTCTAGTAAAAAAACAATTTGTAGAACAGTAGTAGAGCGGCGG 1380
Qy 1539 ATTGTTTTAAATTCATGAAGAATGGGGCAACAACCAAGTGCATCGATCACTGCTTG 1598
Db 1381 ATTGTTTTAAATTCATGAAGACTGGGGCAACAATTCCAAGTGCATCGATCACTGCTTG 1440
Qy 1599 AGCGTGGCAGATGAATAGATGTGCAAGTTTGTATCCACCGATACAGTCAATGAGGCA 1658
Db 1441 AGCGTAGCAGATGAATAGATGTGCAAGTTTGTATACACACCGATACAGTCAATGAGGCA 1500
Qy 1659 GGTATGTAGATGACACCTCAATGAATGAACGCGCGGCATCCATCGCTTACCACATT 1718
Db 1501 GGTATGTAGATGACACCTCAATGAATGAACGCGCGGCATCCATCGCTTACCACATT 1560
Qy 1719 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGGCAGGAGCTCAATATT 1778
Db 1561 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGGCAGGAGCTCAATATT 1620
Qy 1779 CTACCTCTCCACACCCCACTATTCCTATACATTAATACGGTTGCAGAACACTTA 1838
Db 1621 CTACCTCTCCACACCCCACTATTCCTATACATTAATACGGTTGCAGAACACTTA 1680
Qy 1839 GACATGCTCATGACATGACACCATCAGCAACCAACCATCCGCGAGGATTTACAATTTCT 1898
Db 1681 GACATGCTTATGACCTGACACCATCAGCAACCAACCATCCGCGAGGATTTCCAATTTCT 1740
Qy 1899 CAAAGCGTATCCGCGCGGCTCTATCGCGGTGAAGATGTCTCCATGATATGGGTGTG 1958
Db 1741 CAAAGCGTATCCGCGCGGCTCTATCGCGGTGAAGATGTCTCCATGATATCGGTGTG 1800
Qy 1959 ATCGGATGACAAGCTCGGATTCGCAAGCAATGGCGGTGCAGGCGAAGTGAATTCCTCGA 2018
Db 1801 ATCGGATGACAAGTTCGGATTCGCAAGCAATGGCGCGCGCTGGGGAAGTGAATTCCTAGA 1860
Qy 2019 ACTTGGCAGCTCGGATGAAGTAAATAAAGAAATTTGGTGAAGTCTCTGAAGATGCGAAA 2078
Db 1861 ACTTGGCAACTGCGAGAAGAATAAAGAAATTTGGTGAAGTCTCTGAAGATGCGTGA 1920
Qy 2079 GATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACATATCAACCCCGCTTTG 2138
Db 1921 GATAATGACAACTTCGGATCAACCGCTATATCTCAAAATACACCAATTAATCCCGCTTTG 1980
Qy 2139 ACCCAACCGCTGAGCAGATATACGGCTCTGTGGAAGAGGGCAAGATCGCGACTTGGTG 2198
Db 1981 ACCCAACCGCTGAGCAGATATACGGCTCTGTGGAAGAGGGCAAGATCGCGACTTGGTG 2040
Qy 2199 GTGTGGAATCCTGCGCTTTTGGCGTAAACCCAAATCGTGATCAAAAGCGGTATGTTG 2258
Db 2041 GTGTGGAATCCTGCGCTTTTGGCGTAAACCCAAATCGTGATCAAAAGCGGTATGTTG 2100
Qy 2259 GTCTTCTGAAATGGCGGATTTCTAACCGCTCTGTGCCACTTCCCAACCGGTATTATTAC 2318
Db 2101 GTGTTCTCTGAAATGGCGGATTTCTAATGCGCTCTGTGCCACTCTCTCAGCGGTTTATTAC 2160
Qy 2319 CGCGAAATGTTTGGGCATCACGGCAAGGCGAAATTTGACACAGCATCACTTTTGTTC 2378
Db 2161 CGCGAAATGTTTGGGCATCACGGCAAGGCGAAATTTGACACAGCATCACTTTTGTTC 2220
Qy 2379 AAAGTCGCTATGAAATGGCGTGAAGAAAGAGCTGGGCTTAGCGCGCAAGTCTTACGG 2438
Db 2221 AAAGTCGCTATGAAATGGTGTGAAGAAAGAACTAGGTTTAGAGCGCAAGGTGCTCCCC 2280

Qy 2439 GTCAAAACTGCGTAACATCAACGAAGAAGACTTCAAGTTCAACGACAAAACGGCAAAA 2498
Db 2281 GTGAAAAACTGCGTAAACATCAACGAAGAAGACTTCAAGTTCAACGACAAAACGGCAAAA 2340
Qy 2499 ATCACCGTCGATCCGAAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2558
Db 2341 ATCACCGTCGATCCGAAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2400
Qy 2559 CCCACTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTTAGGCACAAT 2610
Db 2401 CCCACTCTGAAGTGCCTCTAGCCCAACGCTACACTTTCTTAGGCATAAT 2452

RESULT 4
AX356695
LOCUS AX356695 2452 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 13 from Patent EP1176192.
ACCESSION AX356695
VERSION AX356695.1 GI:18674032
KEYWORDS
SOURCE
ORGANISM
Helicobacter felis
Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.

REFERENCE
1
AUTHORS Kusters, J.G. and Cottoli, G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: EP 1176192-A 13 30-JAN-2002;
Akzo Nobel N.V. (NL)
FEATURES
Location/Qualifiers
1..2452
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
48...728
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23263.1"
/db_xref="GI:18674033"
/translation="MKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIM
DEARRKKTVAELMECHMFLKKDRVMPGVMNPDVGVEATFPDGTKLVTWNPLEP
DEHFKAGVFCGDKDI ELNAGKRYTELVNTEGKSLHVGSHFHF FEANKLKFRE
KAYGRLOIPSGNTLRIGAGTRKVLQIPLGGSKVIGMNLVNNLIADRRHKHAKLDK
AKSHGFIR"
739..2445
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23264.1"
/db_xref="GI:18674034"
/translation="MKMKQEVNTYPTTGDKVRLGDTDLNAEVEHDYTTVEELKF
GAGTIRGGMGQSNSPDENTLIVITNAMIIDYTGILKADIGIKNGIKHIGIKRAGND
MDGVSPHMVVGTEALAGEMIITAGIDSHTFPLSPQPTTALANGVTTFPGGT
QPGFTNATTTTPGKNLHRLMRAAEYSMNQFLGKNSKSKQIIVEGEAGAIQFK
LHEDGTTPSAIDHCLSVADYVQVCHITDVTNAGYVDDTLNMGRAIYAHYIEG
AGGHSPIVITMAGEVNILPSSTTPTIPVTINTVAHLDMLTCHLDRKIRLEDLQFS
QSRIRPGSIAEDVLHDIGVIAMTSSDSQAMGRAGEVIRPTWQTDKNKKEPKLPED
GADNDFRIKYSKITINPALTGVSEYIGVEBKJADLVVWNPFFGVKPKIVIK
GMVYFBSMGDSNASVPTPPVYREMFHGHGKAFDTSITTFVSKVAYENGVBKJLGL
ERKVLVPVKNCRNITKKDFKENDKATKITVDPKTFEYFVDGKLTCKTSKPTSEVPLAQRYT
FF"

ORIGIN

Query Match 77.3%; Score 2228; DB 2; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 159 AGGACTAATCCCATTCGCTTTTAAATTTTAAACAAGGAGTAATAGGTGAACACTCACAC 218
Db 1 AGGACTAATCCCATTCGCTTTTAAATTTTAAACAAGGAGTAATAGGTGAACACTCACAC 60

QY 219 CCAAAGAGCAAGAAAGTTCTTGTATATATATCGGGCGAAGTGGCTAGAAAGCGCAAG 278
DB 61 CCAAAGAGCAAGAAAGTTCTTGTATATATATCGGGCGAAGTGGCTAGAAAGCGCAAG 120
QY 279 CAGAGGGCTTAAAGCTCAACCAACCGGAGCAATGCTTACATTAAGTGGCCCATATATGG 338
DB 121 CAGAGGGCTTAAAGCTCAACCAACCGGAGCAATGCTTACATTAAGTGGCCCATATATGG 180
QY 339 ACGAAGCGCGCGTGGGAAAGAAACCGTTGGCCAGCTTATGGAGAGTGCATGCACTTTT 398
DB 181 ACGAGGGCGCGTGGGCAAGAAACCGTTGGGAACTTATGGAGAGTGTATGCACTTTT 240
QY 399 TGAAGAGAGTGAAGTAATGCCCGGGTGGTATATGTTCCCGATCTAGGTGTAGAG 458
DB 241 TGAAGAGAGTGAAGTATGCCCGGGTGGGAAATATGGTCCCTGATTTGGGCGTGGAG 300
QY 459 CCACCTTTCCTGATGGTACCAAACTCTGTAAGTGTGAATTTGGCCCATCAAGACAGATGAGC 518
DB 301 CCACCTTTCCTGATGGGACCAAACTCTGTAAGTGTGAATTTGGCCCATCAAGACCTGATGAC 360
QY 519 ACTTCAAAGCGGGGAAAGTGAATTTGGTTGCGATAAGACATCGAGCTCAATCGAGCA 578
DB 361 ACTTTAAGCGGGTGAAGTGAATTTGGTGTGATAAAGACATTTGAACTCAACGCAAGTA 420
QY 579 AAGAGTAAACCGAACTTGAGGTTACTAATGAGGGCCCTAAATCCTTGCATGTGGTAGCC 638
DB 421 AGGAAGTTACCGAACTAGAAAGTTACTAAGCAAGGACCTAAATCCTTGCATGTGGTAGCC 480
QY 639 ATTTCCACTTCTTTGAAGCTTAAAGGACCTTAAATTTGCAATCGTGAAGAAAGCCCTATGSCA 698
DB 481 ATTTCCACTTCTTTGAAGCCCAAGCAATTTGAATTTGCAATTCGATCGGGAAGAAAGCCCTATGSCA 540
QY 699 AAGCCCTAGATATTTCCCTCTGGCAACACGCTACGATTTGGGGCGAGCAACACCGCAAG 758
DB 541 AAGCCCTAGATATTTCCCTCTGGCAACACATACGCAATTTGGGGCGAGCAACACCGTAAAG 600
QY 759 TGCAGTTGATTTCTTGTGGTGGCAGTAAAGAAAGTGTGGCATGAACGGGCTTGTGAATA 818
DB 601 TGCAGTTAATTTCCCTCTTGGCGGTAGTAAAGAAAGTGTGGCATGAACGGGCTTGTGAATA 660
QY 819 ACATCGGGATGAACGCATAAACATAAAGCGCTTGACAGGGCGAATCTCACGGATTTA 878
DB 661 ATATTGGGAGCAAGCGCATAAACACAAAGCGCTTAGACAAAGCAAAATCTCACGGATTTA 720
QY 879 TCAAGTAAGGAGACTCCCATGAAATGAAAGAAACAAAGATATGTAATATCTACGGACCC 938
DB 721 TCAAGTAAGGAGACTCCCATGAAATGAAAGAAACAAAGATATGTAATATCTACGGACCC 780
QY 939 ACCAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC 998
DB 781 ACCAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC 840
QY 999 TATACCACTATGGCGNAGAACTTAAATTTGGGCGGGTAAACTATCCGTGAGGGTATG 1058
DB 841 TATACCACTATGGCGNAGAACTTAAATTTGGGCGGGTAAACTATCCGTGAGGGTATG 900
QY 1059 GGTTCAGAGCAATAGCTGATGAAACACCCCTAGATTTAGTTCATCACTAAAGCGCATGAT 1118
DB 901 GGTTCAGAGCAATAGCTGATGAAACACCCCTAGATTTAGTTCATCACTAAAGCGCATGAT 960
QY 1119 ATCGACTACACCGGGATTTACAAAGCGGACATTTGGGATTTAAAGCGCAAAATCCATGGC 1178
DB 961 ATTTGACTACACCGGGATTTACAAAGCGGACATTTGGGATTTAAAGTGGCAAAATCCATGGC 1020
QY 1179 ATTGGCAAGCGAGAAACAAAGGACATGCAAGATGGCGTAAGCCCTCATATGTGGTGGGT 1238
DB 1021 ATTGGCAAGCGAGAAACAAAGGACATGCAAGATGGCGTAAGCCCTCATATGTGGTGGGT 1080
QY 1239 GTGGGACAGAAAGCACTAGCAGGGAGGTATGATTTATACCGTGGGGGATCGATTCA 1298
DB 1081 GTGGGACAGAAAGCACTAGCAGGGAGGTATGATTTATACCGTGGGGGATCGATTCA 1140
QY 1299 CACACCCACTTCTCTTCCCAACAAATTTCCCTACCGCTCTAGCCATGGCGTTACAAACC 1358

DB 1141 CACACCCACTTCTCTTCCCAACAAATTTCCCTACCGCTCTAGCCATGGCGTTACAAACA 1200
QY 1359 ATGTTTGGAGGGCGGACAGGTCCTGTAGATGGCAAGATGGCACTACTATCACTCCGGGC 1418
DB 1201 ATGTTTGGCGGTGGACAGGCCCGGTAGATGGCAAGATGGCACTACTATCACTCCGGGC 1260
QY 1419 AAATGGAACTTGCACCGCATGTTGGCGCGCAGCAGAGAGTATTTCTATGATATGGGCTTT 1478
DB 1261 AAATGGAACTTGCACCGCATGTTGGCGCGCAGCAGAGAGTATTTCTATGATATGGGCTTT 1320
QY 1479 TTGGGCAAGGCAATAGCTCTAGCAAAACAACTTTGTAGAAACAGTGTAGAAAGCGGGCGG 1538
DB 1321 TTGGGCAAGGCAATAGCTCTAGTAAACAACTTTGTAGAAACAGTGTAGAAAGCGGGCGG 1380
QY 1539 ATTGGTTTTAAATGTCATGAAGACTGGGGCAACCAACCAAGTGCATCGATCACTGCTTG 1598
DB 1381 ATTGGTTTTAAATGTCATGAAGACTGGGGCAACCACTCCAAGTGCATCGATCACTGCTTG 1440
QY 1599 AGCGTGGCAGATGAATACGATGTGCAAGTGTGTATCCACCGATACAGTCAATGAGGCA 1658
DB 1441 AGCGTGGCAGATGAATACGATGTGCAAGTGTGTATCCACCGATACAGTCAATGAGGCA 1500
QY 1659 GGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCCATCCATGCTACCACTT 1718
DB 1501 GGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCCATCCATGCTACCACTT 1560
QY 1719 GAGGAGCGGGTGGAGGACACTCACTGATGTTATCACCATGGCAGCGAGCTCAATATTT 1778
DB 1561 GAGGAGCGGGTGGAGGACACTCACTGATGTTATCACCATGGCAGCGAGTGAATATTT 1620
QY 1779 CTACCTCTCCACCAACCCCTACTATCCCTATACCAITAAATACGGTTGCGAGAACTTTA 1838
DB 1621 CTACCTCTCCACCAACCCCTACTATCCCTATACCAITAAATACGGTTGCGAGAACTTTA 1680
QY 1839 GACATGCTCATGACATGCGCACACCTAGACAAACCGCATCCGCGAGGATTTACAAATTTCT 1898
DB 1681 GACATGCTTATGACCTGCGCACACCTAGATAAACCGCATCCGCGAGGATCTCCAATTTCT 1740
QY 1899 CAAAGCGGTATCCGCGCGGCTCTATCCGCGGTGAAGATGTCTCCATGATATGGGTGTG 1958
DB 1741 CAAAGCGGTATCCGCGCGGCTCTATCCGCGGTGAAGATGTCTCCATGATATCGGTGTG 1800
QY 1959 ATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGGCAAGTATCTCTCGA 2018
DB 1801 ATCGCGATGACAAGTTCGATTTGCAAGCAATGGGCGGCTGGGCAAGTATCTCTAGA 1860
QY 2019 ACTTGGCAGACTGCGGATGAAGATAAAGAAAGATTTGGTAAAGCTTCTCGAAGTGCATA 2078
DB 1861 ACTTGGCAAACTGCGAGATATCGGCTCTGTGGAAGAGGGCAAGATCCCGCACTTGGTG 1920
QY 2079 GATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACATATCAACCCCGCTTTG 2138
DB 1921 GATATGACAACTTCCGCAATCAAGCGCTATATCTCCAAATACCAATTAATCCCGCTTTG 1980
QY 2139 ACCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATCCCGCACTTGGTG 2198
DB 1981 ACCCATGGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATCCCGCACTTGGTG 2040
QY 2199 GTGTGGAAATCTGCTTTTTTTGGCGTAAAAACCCAAATCTGTATCAAGCGGTATGGTG 2258
DB 2041 GTGTGGAAATCTGCTTTTTTTGGCGTAAAAACCCAAATCTGTATCAAGCGGTATGGTG 2100
QY 2259 GTCTTCTCTGAATGGGCGATTTCTAACCGCTGTGCCACTCCCAACCGGTTTATATAC 2318
DB 2101 GTGTCTCTGAATGGGCGATTTCTAATCGCTGTGTGCCACTCTCAGCGGTTTATATAC 2160
QY 2319 CGGAAATGTTTGGGCATCAGCGCAAGCGGAAATTTTGACACAGCATCACTTTGTTTTC 2378
DB 2161 CGGAAATGTTTGGGCATCAGCGCAAGCGGAAATTTTGACACAGCATCACTTTGTTTTC 2220
QY 2379 AAAGTTCGCTATGAAATGGCGTGAAGAAAGAGTGGGCTTAGAGCGCAAGTTCCTACCG 2438

2221	AAAGTCGCCTATGAAATAATGGTGTGAAAGAAAAAATAGTGTAGAGCGCAAGGTGCTCCCC	2288
2439	GTCAAAAATCGCCGTAAACATCACCAAGAAAGACTTCAAGTTTCAACGACAAAACGCGAAAA	2498
2281	GTGAAAAATCGCGTAACATCACCAAGAGGACTTCAAGTTTCAACGACAAAATGCGAAAA	2340
2499	ATCACCGTCGATCCGAAAAACCTTCGAGGCTTTGTGTAGATGGCAAACTCTGCACCTCTAAA	2558
2341	ATCACCGTCGATCCGAAAAACCTTCGAGGCTTTGTGTAGATGGCAAACTCTGCACCTCTAAA	2400
2559	CCCACTCGCAAGTGCTCTTAGCCAGCGCTACACTTTCTTCTAGGCACAAAT	2610
2401	CCCACTCTGAAGTGCTCTTAGCCCAACGCTACACTTTCTTCTAGGCATAAT	2452
RESULT 5		
BD185303	2405 bp	DNA
LOCUS	linear	PAT 17-JUN-2003
DEFINITION	Helicobacter felis vaccine.	
ACCESSION	BD185303	
VERSION	BD185303.1	GI:31877503
KEYWORDS	JP 2002355054-A/2.	
SOURCE	Helicobacter felis	
ORGANISM	Helicobacter felis	
REFERENCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;	
AUTHORS	Helicobacteraceae; Helicobacter.	
TITLE	1 (bases 1 to 2405)	
JOURNAL	Kusters, J.G. and Cattoli, G.	
	Helicobacter felis vaccine	
	Patent: JP 2002355054-A 2 10-DEC-2002;	
	AKZO NOBEL NV	
COMMENT	OS Helicobacter felis	
	PN JP 2002355054-A/2	
	PD 10-DEC-2002	
	PF 16-JUL-2001 JP 2001214711	
	PR 17-JUL-2000 EP 00202565.8	
	PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI	
	PC	
	C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC	
	A61K39/23	
	PC A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15,	
	PC C12N1/19,	
	PC	
	C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/ PC	
	53,	
	PC	
	G01N33/53, G01N33/566, G01N33/569/(C12N9/80, C12R1/01), (C12Q1/68, PC	
	C12R1/01),	
	PC C12N15/00, C12N5/00, A61K37/02	
	CC Helicobacter felis vaccine	
	PH Key Location/Qualifiers	
	FT CDS (1) (681)	
	FT CDS (692) (2398).	
FEATURES	Location/Qualifiers	
source	1. 2405	
	/organisms="Helicobacter felis"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:214"	
ORIGIN		
Query Match	74.9%;	Score 2158.6; DB 2; Length 2405;
Best Local Similarity	93.6%;	Pred. No. 0;
Matches 2251;	Conservative	0; Mismatches 154; Indels 0; Gaps 0;
Qy	206	GTGAAGTACACCCCAAGCAAGAAAGTTCTTGTATATATATGCGGGCGAAGTGGCT 265
Db	1	GTGAAGTACACCCCAAGCAAGAAAGTTCTTGTATATATATGCGGGCGAAGTGGCT 60
Qy	266	AGAAAGCGAAAGCAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 325
Db	61	AGAAAGCGAAAGCAGAGGGCTTAAAGCTCAATCAACCGAAGCCATTGCTTACATTAGT 120
Qy	326	GCCCATATTATGACGAAGCGCGCTGGGAAAAAATCCGTTGCCAGCTTATGGAAGAG 385

```
Db 1201 CATCACTCGGGCAATGGAATTTGCACCGCATGTTGGCGCGAGCTGAAGAGTATTCTAT 1260
Qy 1466 GAATGTGGCTTTTGGGCAAGCAATAGCTCTAGCAAAAAAACAATTGTAGAACAACT 1525
Db 1261 GAATGTGGCTTTTGGGCAAGCAATAGCTCTAGCAAAAAAACAATTGTAGAACAACT 1320
Qy 1526 AGAAGCGGGCGGATTCGTTTAAATTCATGAAGACTGGGGCAACAACCAAGTGGCAT 1585
Db 1321 AGAAGCGGGCGGATTCGTTTAAATTCATGAAGACTGGGGCAACAACCAAGTGGCAT 1380
Qy 1586 CGATCACTGCTTGAGCGTGGCAGATGAATAGATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTGCTTGAGCGTGGCAGATGAATAGATGTGCAAGTTTGTATCCACACCGATAC 1440
Qy 1646 AGTCAATAGAGCGGTTATGTAGATGACACCTTAAATGCAATGAAGCGGCGCATCCA 1705
Db 1441 GGTCAATAGAGCGGTTATGTAGATGACACCTTAAATGCAATGAAGCGGCGCATCCA 1500
Qy 1706 TGCTACCAATTGAGGAGCGGTGGAGGACACTCACTGATGTTTATCACCATGGCAGG 1765
Db 1501 TGCTACCAATTGAGGAGCGGTGGAGGACACTCACTGATGTTTATCACCATGGCAGG 1560
Qy 1766 CGAGCTCAATATTCACCTCTCCACACACCCCACTATTCCCTATACCATTAATACGCT 1825
Db 1561 CGAGCTCAATATTCACCTCTCCACACACCCCACTATTCCCTATACCATTAATACGCT 1620
Qy 1826 TGCAAGACACTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1885
Db 1621 TGCAAGACACTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Qy 1886 TTTACAAATTTTCAAGCGGTATCGCGCGGCTCTATCGCGGCTGAAGATGCTTCCA 1945
Db 1681 TCTCCAGTTTTCACAGCGGTATCGCGCGGCTCTATCGCGGCTGAAGATGCTTCCA 1740
Qy 1946 TGATATGGGTGATCGGATGATCAAGCTCGGATTCGCAAGCAATGGGGCTGCAGCGGA 2005
Db 1741 TGATATGGGTGATCGGATGATCAAGCTCGGATTCGCAAGCAATGGGGCTGCAGCGGA 1800
Qy 2006 AGTGATTTCTCGAATTTGGCAGACTGGCGATAAGATAAAGAAATTTGTAGCTTCC 2065
Db 1801 AGTGATTTCTCGAATTTGGCAGACTGGCGATAAGATAAAGAAATTTGTAGCTTCC 1860
Qy 2066 TGAAGATGGCAAGATAACGATTAATTTCCGATTAAGCGCTACATCTCCAAATACATAT 2125
Db 1861 TGAAGATGGCGAGATAATGACAACTTCGCGATCAACGCTATATCTCCAAATACACAT 1920
Qy 2126 CAACCCGCTTTGACCCACCGCGTGAGAGTATATCGGCTCTGTGCAAGAGGCGAAGAT 2185
Db 1921 TAATCCGCTTTGACCCACCGCGTGAGAGTATATCGGCTCTGTGCAAGAGGCGAAGAT 1980
Qy 2186 CGCGGACTTGGTGTGGAATCTGCTTTTGGCGTAAAAACCCAAATCGTGATCAA 2245
Db 1981 CGCGGACTTGGTGTGGAATCTGCTTTTGGCGTAAAAACCCAAATCGTGATCAA 2040
Qy 2246 AGCGGATTTGGTGTCTCTGAAATGGCGGATTTCAACGCGTCTGTGCGCATCTCCCA 2305
Db 2041 AGCGGATTTGGTGTCTCTGAAATGGCGGATTTCAACGCGTCTGTGCGCATCTCCCA 2100
Qy 2306 ACCGTTTATACCGGCAATGTTTGGGATTCACGGCAAGCGGCAAAATTTGACACCGCAT 2365
Db 2101 GCGGTTTATACCGGCAATGTTTGGGATTCACGGCAAGCGGCAAAATTTGACACCGCAT 2160
Qy 2366 CACTTTGTTTCCAAAGTCGCTATGAAATGCGGTGAAGAAAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTGTTTCCAAAGTCGCTATGAAATGCGGTGAAGAAAGCTTAGAGCG 2220
Qy 2426 CCAAGTTCTACCGGTCAAAAATCGCGTAAACATCAACCAAGAAAGCTTCAAGTTCAACGA 2485
Db 2221 CAAGTGTCTACCGGTCAAAAATCGCGTAAACATCAACCAAGAAAGCTTCAAGTTCAACGA 2280
Qy 2486 CAAACGGCAAAAATCACCGTGCATCCGAAACCTTCGAGGCTTTGTAGATGCGCAACT 2545
Db 2281 CAAGCGGCGCATATCACTGTGCATCTTAAACCTTCGAGGCTTTGTAGATGCGCAACT 2340
```

```
Qy 2546 CTGCACTCTAAACCCACCTCGCAAGTGCCTTAGCCAGCGCTACACTTTCTTAGGC 2605
Db 2341 CTGCACTCTAAACCCCGCTCTGAAGTGCCTTAGCCCAACGCTACACTTTCTTAGGC 2400
Qy 2606 ACAAT 2610
Db 2401 ACAAT 2405

RESULT 6
AX356686
LOCUS AX356686 2405 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 4 from Patent EP1176192.
ACCESSION AX356686
VERSION AX356686.1 GI:18674023
KEYWORDS
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE
1 Kusters, J. G. and Cattoli, G.
AUTHORS Helicobacter felis vaccine
TITLE Patent: EP 1176192-A 4 30-JAN-2002;
JOURNAL Akzo Nobel N.V. (NL)
FEATURES
Location/Qualifiers
source
1. 2405
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
1. 681
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD3257.1"
/db_xref="GI:18674024"
/translation="MKLTPKEQKFLYYAGEVARKKRAEGLKLNQPEAIAYISAHIM
DEARRGKTVAEALMEECMHFLKDEKDEVMGVMVDPDLGVEATPDGKTLVITVWPIEP
DEHFKACDVGKCDKIDELNAGKEVTELEVNKGPKSLHVGSHHFFETFNKALKPDE
KAVGKRLDIPSGNTRIGAGQTRKQLIPLGSSKKVIGMGLVNNIADERHKKHALDK
AKSHGPIK"
692..2398
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD3258.1"
/db_xref="GI:18674025"
/translation="MKMKQEVNTYGYPTTGDVKRLGDTDLWAEVYEHDTTYTYBELKF
GAGKTIREGMQSPDENTLIDLVTNAMIIDYTGIVKADIGIKNGKIHGIGKGNKD
MGQVSPHMVVGVTALAGEGMIITAGIDSHTHFLSPQOPPTALANGVITMFGGT
GPVDGTVATTITPGKWLHBLRAEYSNMVFLGKNSKSKOLVQVEAGAIGFK
LHEDWGTTPSAIDHCLSVADBYDVQVCIHTDVTNEAGYVDITLNNMNGRAIHAYHIEG
AGHGSDDVITMAGELNLPSTTPTIPTYITVAHLDMLMTCHLLDKRIHEDIQFS
QSRIRPQIAAEDVLHDIGVITAMTSSDQAMGRAGEVITRTWQTDKNKKEFKLPED
GANDNPKRIYSKYITNPALTHGVSEYIGSVGEKIADLVWNPAPFGVVKPIVIK
GGMVFSMGSDNASVPTPOPVYVREMFHGHGKAFDTSITFVSKVAVENGVEKGLGL
ERKVLVYKNCNITKDPKFNKTAHITVDPKTEFVFDGKLTCKSPASEVPLAQRYT
FP"
```

ORIGIN

Query Match 74.9%; Score 2158.6; DB 2; Length 2405;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

```
Qy 206 GTGAACTCACCCAAAGAGCAAGAAAGTCTTGTATATATTCGCGGCGAAGTGGCT 265
Db 1 GTGAACTCACCCAAAGAGCAAGAAAGTCTTGTATATATTCGCGGCGAAGTGGCT 60
Qy 266 AGAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 325
Db 61 AGAAGCGCAAGCAGAGGCTTAAAGCTCAATCAACCGAAGCCATTGCTTACATTAGT 120
```


QY 326 GCCCATATTATGGAGAGAGCGCGCGTGGAAAAAACCAGTTCGCCAGCTTATGGAAGAG 385
Db 121 GCCCATATTATGGAGAGAGCGCGCGTGGAAAAAACCAGTTCGCCAGCTTATGGAAGAA 180
QY 386 TGCATGCACTTTTGA AAAAAGATGAAGTAATGCCGGGTGGGTAAATATGTTCCCGAT 445
Db 181 TGTATGCACATTTTGA AAAAAGATGAAGTAATGCCGGGTGGGGAATATGTTCCCTGAT 240
QY 446 CTAGGTGTAGAAGCCACCTTTCTCTGATGGTACGAAACTTGTAACTGTGAAATTGGCCCATC 505
Db 241 TTGGGCGGTAGAAGCCACTTTCCCGATGGCACCAAACTCGTAAACCGTGAATTGGCCCAT 300
QY 506 GAACCCAGATGAGCACTTAAAGCGGGCGAAGTGAATTTGGTTGGATGAAGACATGAG 565
Db 301 GAACTGTATGAACACTTTAAAGCGGGTGAAGTGAATTTGGCTGTGATAAAGACATTCGAG 360
QY 566 CTCAATCAGGCAAGAGTAACCGAAGTTCAGGTTACTAATGAAGGCGCTAAATCTCTTG 625
Db 361 CTCAACGGGTAGAAGATTACCGAGCTTGAAGTTACCAAGAGGACCTAAATCTCTTG 420
QY 626 CATGTGGGTAGCCATTTCACCTTTCTTGAAGCTAAACAGGCACTAAATTCGATCGTGA 685
Db 421 CATGTGGGTAGCCATTTCACCTTTCTTGAACCAACAGGCACTTGAATTCGATCGGGA 480
QY 686 AAAGCCTATGCAAAACGCTTAGATATTCCTCTCTGGCAACACGCTACGCAATGGGGCAGGA 745
Db 481 AAAGCCTATGCAAAACGCTTAGATATTCCTCTCTGGCAACAGCTACGCAATGGGGCAGGA 540
QY 746 CAACCCGCAAGTGCAGTTGATTCCTCTCTGGTGCAGTAAAGTGAATGGCAATGAAC 805
Db 541 CAACCCGCTAAGTGCAGTTAATCCCTCTTGGCGGTAGTAAAGTGAATGGCAATGAAC 600
QY 806 GGGCTGTGTAATAACATCGCGGATGAACGCATATAACATAAGCCCTTGAACAAGCCGAAA 865
Db 501 GGGCTGTGTAATAATTTGCGGACGAAGCCATATAACACAAGCACTAGACAGGCAAAA 660
QY 866 TCTCAGCGATTTATCAAGTAAGAGAGATCCCATGAAATGAAAAACAAGAAATATGTAAA 925
Db 661 TCTCAGCGATTCATCAAGTAAGAGAGATCCCATGAAATGAAAAACAAGAGTATGTAAA 720
QY 926 TACCTACGGACCCACCAAGGCGATAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGA 985
Db 721 CACCTACGGACCCACCAAGGCGATAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGA 780
QY 986 AGTAGAACATCACTATACCATTTAGGCGAAGAACCTTAAATTTGGCGGGTAAAACTAT 1045
Db 781 AGTAGAACATCACTATACCATTTAGGCGAAGAGCTCAAAATTTGGCGGGTAAAACTAT 840
QY 1046 CCGTGAGGGTATGGGTACAGAGCAATAGCCCTGATGAAAAACCCCTAGATTTTAGTCATCAC 1105
Db 841 CCGTGAGGGTATGGGTACAGAGCAATAGCCAGATGAAAAACCCCTTAGATTTTAGTCATCAC 900
QY 1106 TAACCGGATGATTTATGCACTACACCGGATTTAACAAGCGGACATTTGGGAATTA AAAACGG 1165
Db 901 CAACCGGATGATTTATGCACTACACCGGATTTTATAAGCCGACATTTGGTATTA AAAAATGG 960
QY 1166 CAABAATCATCGCATTTGCAAGGAGGAAACAAGCACTGAAGATGCGTAAAGCCCTCA 1225
Db 961 CAABAATCATCGCATTTGCAAGGAGGAAACAAGCACTGAAGATGCGTAAAGCCCTCA 1020
QY 1226 TATGCTCGTGGGTGGGACAGAGCACTAGCAGGGGAAGGTATGATTTATACCGCTGG 1285
Db 1021 TATGCTCGTGGGTGGGACAGAGCACTAGCAGGGGAAGGTATGATTTATACCGCTGG 1080
QY 1286 GGGAAATCGATTTACACACCCACTTCTCTTTTCCCAACAATTTCCCTACCGCTCTAGCCAA 1345
Db 1081 GGGGATCGATTTGCGACACCCACTTCTCTCTCCCAACAATTTCCCTACCGCTCTAGCCAA 1140
QY 1346 TGGCGTTACCAACCATTTTGGAGGGGCGACAGGCTCTGTAGTGCAGGATGCGACTAC 1405
Db 1141 TGGGTGTTACCAACCATTTTGGAGGGGCGACAGGCTCTGTAGTGCAGGATGCGACTAC 1200
QY 1406 TATCACTCCGGGCAAAATGGAACCTTGCAACCGCATGTTGGCGGAGCAGAGATTTCTAT 1465

Db 1201 CATCACTCCGGCAAAATGGAACCTTGCAACCGCATGTTGCGCGCAGCTGGAAGAGTATTTCTAT 1260
QY 1466 GAAATGTGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAACAAAGT 1525
Db 1261 GAAATGTGGCTTTTGGGCAAAAGGCAATAGCTCTAGTAAAAAACAACCTCGTAGAACAAAGT 1320
QY 1526 AGAAGCGGGCGGATTTGGTTTAAATTCATGAAGCTCGGGGCAACCAACCAAGTCCGAT 1585
Db 1321 AGAAGCGGGCGGATTTGGCTTTAAATTCATGAAGCTGGGGCAACCAACCAAGTCCGAT 1380
QY 1586 CGATCACTGCTTGAGCGTGGCAGATGAATACCATGTGCAAGTTTGTATCCACCCGATAC 1645
Db 1381 CGATCACTGCTTGAGCGTGGCAGATGAATACCATGTGCAAGTTTGTATCCACCCGATAC 1440
QY 1646 AGTCAATGAGGCGAGTATGTAGATGAACCCCTAAATGCAATGAACGGGCGGCCATCCA 1705
Db 1441 GGTCAATGAGGCGAGTATGTAGATGAACCCCTAAATGCAATGAACGGGCGGCCATCCA 1500
QY 1706 TGCCTACCACTTTAGAGGAGCGGGTGGAGGACCTCACCTGATTTTATCACCATGACAGG 1765
Db 1501 TGCCTACCACTTTAGAGGAGCGGGCGGAGGACCTCACCTGATTTTATCACCATGCGAGG 1560
QY 1766 CGAGCTCAATATTCTACCTCTCTCACCCGCCACTATTCCTTATACCATTAATACGGT 1825
Db 1561 CGAGCTCAATATTCTACCTCTCTCACCCGCCACTATTCCTTATACCATTAATACGGT 1620
QY 1826 TGCAGAACCACTTTAGACATGCTCATGACATGCCACCACTAGACAAAACGATCCGCGAGGA 1885
Db 1621 TGCAGAACCACTTTAGACATGCTCATGACCTGCCACCACTAGACAAAACGATCCGCGAGGA 1680
QY 1886 TTTTCAAAATTTTCTCAAAAGCGGTATCCGCCCGCTCTATCGCGGCTGAAGATGTGTCTCA 1945
Db 1681 TCTCCAGTTTTTCCCAAGCGGTATCCGCCCGCTCTATTTGCCGCTGAAGATGTGTCTCA 1740
QY 1946 TGATATGGTGTGATCGCGATGCAAGCTCGGATTTGCGAAGCAATGCGGCGTGCAGCGGA 2005
Db 1741 TGATATTTGGCGTGTATCGCGATGCAAGCTCGGATTTGCGAAGCAATGCGGCGCTCGGGA 1800
QY 2006 AGTGAATCTCTCGAATCTTGCAGACTGCGGATGAAGATAAAGAAATTTGGTAAAGCTTCC 2065
Db 1801 AGTGAATCTCTAGAACTTGGCAAACTGCGAAGAAATAAAGAAATTTGGTAAAGCTTCC 1860
QY 2066 TGAAGATGCGCAAGATAACGATAATTTTCGCAATTAAGCGCTACATCTCCAAAATACACTAT 2125
Db 1861 TGAAGATGCTGACAGATAATGACAACTTCGCGATCAAAACCTATATCTCCAAAATACACTAT 1920
QY 2126 CAAACCCGCTTTGACCCACCGGCTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT 2185
Db 1921 TAAATCCCGCTTTGACCCCATGGCGTGAAGCGATATATCGGCTCTGTGGAAGAGGCGCAAGAT 1980
QY 2186 CGCGCACTTGGTGTGTAATCTTCCCTTTTGGCGTAAACCCCAAAATCGTGATCAA 2245
Db 1981 CGCGCACTTGGTGTGTAATCTTCCCTTTTGGTGTAAACCCCAAAATCGTGATCAA 2040
QY 2246 AGCGCGTATGGTGTCTTCTCTGAAATGGCGCATCTAAACGGGCTCTGTGCGCACTCCCA 2305
Db 2041 AGCGCGTATGGTGTCTTCTCTGAAATGGCGCATCTAAACGGGCTCTGTGCGCACTCCCA 2100
QY 2306 ACCGGTTATTTACCCGCAAAATTTTGGGCAATCAAGCGCAAGGCGGAAATTTGACACCGAGCAT 2365
Db 2101 GCGCGTTTATTTACCGCGAAATGTTTGGGCAATCAAGCGCAAGGCGGAAATTTGACACCGAGCAT 2160
QY 2366 CACTTTTGTTCGAAGTGCCTATGAAATGCGGTGAAGAAAGCTGGCTTAGAGCG 2425
Db 2161 CACTTTTGTTCGAAGTGCCTATGAAATGCGGTGAAGAAAGCTAGGCTTAGAGCG 2220
QY 2426 CCAAGTCTTACCGGTCAAAAACCTGCGTAAACCTCACAGAAAGACTTCAAGTTTCAACGA 2485
Db 2221 CAAAGTCTTACCGGTGA AAAAAGCTGCGCAACATCACTAAGAAAGACTTCAAAATTCACAA 2280
QY 2486 CAAAAACGCGCAAAATCAACCGTGCATCCGAAAAACCTTTCGAGGCTTTTGTAGATGGCAAACT 2545


```

Db      2281 CAAGACGGCGCATATCACTGTGATCTTAAACCTTCGAGGTCTTTGTAGATGGCAACT 2340
Qy      2546 CTGCACCTCTAAACCCACCTCGAAGTGCCTCTAGCCAGCGCTACACTTCTCTAGGC 2605
Db      2341 CTGCACCTCTAAACCCCGCTCTGAAGTGCCTCTAGCCAGCGCTACACTTCTCTAGGC 2400
Qy      2606 ACAAT 2610
Db      2401 ACAAT 2405

RESULT 7
BD185305
LOCUS   BD185305                2407 bp      DNA      linear      PAT 17-JUN-2003
DEFINITION Helicobacter felis vaccine.
ACCESSION BD185305
VERSION   BD185305.1 GI:31877505
KEYWORDS  JP 2002355054-A/4.
SOURCE   Helicobacter felis
ORGANISM Helicobacter felis
          Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
          Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 2407)
AUTHORS   Kusters,J.G. and Cattoli,G.
TITLE     Helicobacter felis vaccine
JOURNAL   Patent: JP 2002355054-A 4 10-DEC-2002;
          AKZO NOBEL NV
COMMENT   OS Helicobacter felis
          PN JP 2002355054-A/4
          PD 10-DEC-2002
          PF 16-JUL-2001 JP 2001214711
          PR 17-JUL-2000 EP 00202565.8
          PI JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI
          PC
          C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC
          A61K39/23,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,
          PC C12N1/19,
          PC
          C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC
          53,
          PC
          G01N33/53,G01N33/566,G01N33/569//(C12N9/80,C12R1/01),(C12Q1/68, PC
          C12R1/01),
          PC C12N15/00,C12N5/00,A61K37/02
          CC Helicobacter felis vaccine
          FH Key Location/Qualifiers
          FT CDS (2)..(682)
          FT CDS (693)..(2399).
FEATURES
source   1..2407
          /organism="Helicobacter felis"
          /mol_type="genomic DNA"
          /db_xref="taxon:214"

ORIGIN
Query Match 74.9%; Score 2158.6; DB 2; Length 2407;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy      206 GTGAAGCTCAACCCAAAGAGCAAGAAAGTCTTGTATATATATGCGGCGAAGTGGCT 265
Db      2 GTGAAGCTCAACCCAAAGAGCAAGAAAGTCTTGTATATATATGCGGCGAAGTGGCT 61
Qy      266 AGAAGCGCAAGACGAGGCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT 325
Db      62 AGAAGCGCAAGACGAGGCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT 121
Qy      326 GCCCATATATGACGACGCGCGTGGAAAAAAGCCGTTGCCAGCTTATGGAAGAG 385
Db      122 GCCCATATATGACGAGGCGCGCGTGGAAAAAAGCCGTTGCCAGCTTATGGAAGAG 181
Qy      386 TGCAATGCACTTTTGTAAAAAAGATGAAGTAATCCCGGGGTGGTAAATATATGTTCCCGAT 445

```

```

Db      182 TGTATGCACTTTTGAANAAGACGAGGTGATCCCGGTGTGGGAATATGGTCCCTGAT 241
Qy      446 CTAGGTGTAGAAAGCCACCTTTCTCTGATGGTAGAAACTTTGTAACTGTGTAATGGCCCATC 505
Db      242 TTAGGCGTGAAGACTACTTTTCCCGATGGCACCAACTCGTAACCGTGAATTTGGCCCATC 301
Qy      506 GAAACAGATGAGACCTTCAAAGCGGCGGAAGTGAATTTGGTTCGATAAAGACATCGAG 565
Db      302 GAAACCGATGAACACTTCAAAGCGGCGGAAGTCAAATTTGGCTGTGTATAAAGACATTTGAA 361
Qy      566 CTCATGTCAGGCAAGAACTTAACCTTGAAGTGTACTAAATGAAGGCGCTAAATCCTTG 625
Db      362 CTCACAGCAGGTGAAGAGTTTACCGAACTAGAAGTTTACCACGAAGACCTAAATCCTTG 421
Qy      626 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCTTAAACAGGCCTAAATTTTCGATCGTGA 685
Db      422 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCCCAAGGCATTGAATTCGATCGGAA 481
Qy      686 AAGCCTATGGCAAAAGCGCTAGATATTCCTCTGGCAACACGCTACGATTTGGGGCAGGA 745
Db      482 AAGCCTATGGCAAAAGCGCTAGATATTCCTCTGGCAACACGCTACGATTTGGGGCAGGA 541
Qy      746 CAACCCGCAAGTGCAGTTGATTTCTTGTGGCAGTAAAGAGTGAATTTGGCATGAAC 805
Db      542 CAACCCGTAAGTGCAGTTAATTCCTCTTGGCGGCGAGTAAAGAGTGAATTTGGCATGAAC 601
Qy      806 GGGCTTGTGAATAACATCGCGGATGAACGCCATAAATGAAGCGCTTCAAGAGGCGAA 865
Db      602 GGGCTTGTGAATAATATTTGCAGATGAACGCCATAAATGAAGCGCTTGAAGAGCAAA 661
Qy      866 TCTCAGGATTTATCAAGTAAGGAGACTCCCATGAAATGAAGAGTGAATTTGATGATA 925
Db      662 TCTCAGGATTTATCAAAATGAAGGAGACTCCCATGAAATGAAGAGTGAATTTGATA 721
Qy      926 TACTTACGCAACCCACCAAGGCGATAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAGA 985
Db      722 TACTTACGCAACCCACCAAGGCGATAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAGA 781
Qy      986 AGTAGAATGACATATACCATCTATGGCGGAAGAACTTAAATTTGGCGGGTAAAACTAT 1045
Db      782 AGTAGAATGACATATACCATCTATGGCGGAAGAACTTAAATTTGGCGGGTAAAACTAT 841
Qy      1046 CCGTGAGGCTATGGTTCAGAGCAATAGCCCTGATGAAACACCCCTAGATTTAGTCAATCAC 1105
Db      842 CCGTGAGGCTATGGTTCAGAGCAATAGTCCAGATGAAACACCCCTAGATTTAGTCAATCAC 901
Qy      1106 TAAACCGATGATTATCGACTACACCGGATTTTAAAGCCGACATTTGGGATTTAAACCGG 1165
Db      902 CAACCGATGATTATGACTACACCGGATTTTAAAGCCGACATTTGGGATTTAAACCGG 961
Qy      1166 CAAATCCATGGCATTTGCAAGGCGGAAGAAACAGGACATGCAAGATGGCGTAAAGCCCTCA 1225
Db      962 CAAATCCATGGCATTTGCAAGGCGGAAGAAACAGGACATGCAAGATGGCGTAAAGCCCTCA 1021
Qy      1226 TATGTCGTGGGTGGGCAAGAGCACTAGCAGGGAAGGTATGATTTATACCGCTGG 1285
Db      1022 TATGTCGTGGGTGGGCAAGAGCATTTAGCAGGGAAGGTATGATTTATACCGCTGG 1081
Qy      1286 GGGAAATCGATTCAACACCCACTTCTCTTCCCAACAAATTTCCCTACCGCTCTAGCCAA 1345
Db      1082 GGGAAATCGATTCAACACCCACTTCTCTTCCCAACAAATTTCCCTACCGCTCTAGCCAA 1141
Qy      1346 TGGCGTTACACCATGTTTGGAGGCGGCAAGGCTCTGTAGATGCGAATGCGACTAC 1405
Db      1142 TGGCGTTACACCATGTTTGGAGGCGGCAAGGCTCTGTAGATGCGAATGCGACTAC 1201
Qy      1406 TATCACTCCGGGCAAAATGGAATCTTGACCGCATTTGGCGCAGCAGAGAGTATTCTAT 1465
Db      1202 CATCACTCCGGGCAAAATGGAATCTTGACCGCATTTGGCGCAGCAGAGAGTATTCTAT 1261
Qy      1466 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTTAGCAAAAAACAACTTTGTAGAACAACT 1525

```

Qy	2606	ACAATG 2611
Db	2402	NCAATG 2407
RESULT 8		
AX356692		
LOCUS	AX356692	2407 bp DNA linear PAT 13-FEB-2002
DEFINITION	Sequence 10 from Patent EP1176192.	
ACCESSION	AX356692	
VERSION	AX356692.1	GI:18674029
KEYWORDS	Helicobacter felis	
SOURCE	Helicobacter felis	
ORGANISM	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Helicobacter.	
REFERENCE	1	
AUTHORS	Kusters,J.G. and Cattoli,G.	
TITLE	Helicobacter felis vaccine	
JOURNAL	Patent: Ep 1176192-A 10 30-JAN-2002;	
FEATURES	Akzo Nobel N.V. (NL)	
source	Location/Qualifiers	
CDS	1..2407	
	/organism="Helicobacter felis"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:214"	
	2..682	
	/note="unnamed protein product"	
	/codon_start=1	
	/transl_table=11	
	/protein_id="CAD23261.1"	
	/db_xref="GI:18674030"	
	/translation="MKLTPEQSKFLLYAGEVARKKAEGLKLNOPEALAYISAHIM DEARGGKTVAELMERCHEFLKDEVMGVGNVPDLGVETPFGDKGLIVNWPTEIP DEHPKAGEVPGDKDIENAGKEVTELEVTNEGPKSLHVGSHFHPPEANKALKKFDRER KAYGRUDIPSGNTLRIGAGTRKVQLIPLGGSSKKVI GMLGNLVNNIADERRHKHALEK AKSHGFIK"	
CDS	693..2399	
	/note="unnamed protein product"	
	/codon_start=1	
	/transl_table=11	
	/protein_id="CAD23262.1"	
	/db_xref="GI:18674031"	
	/translation="MMKKQEVYNTYGPTTGDKRVLDGTDLWAEVEHDYTTYGEELKF GAGKTIREGQGSNPENTLDLVI TNAMIIDVTGIVKADIGIKNGKHIGIKAGNDT MQDVSFMNVVGVGTALACEGMII TAGGIDSHTEFLSPQOPTALANGVTTMGGGT GPVDGNTAII PGKNWLHRLRAEYSNNVGLFGKNSSSKKQLVEQIEAGAIGFK LEDHWGTPSAIDHCLSVADEYDVQCITDTVNEAGYVDDTTLNAMNGRAIHAYHLEG AGHSDDPSPI TMAGELNI LPSSTPTPIYTINTVAHLDMLTCHHLDKR IREDLPQS QSR LRPGSI AAEVLHDIGVIA MTSDSOAMRAGEVI PRWTOTADKNKKEFGKLPED SANDNRIRKYSKITINDPAL THGVSEVIGSVEEGKIADLVVNPNPAFCGVKPKVILK GMVVFSEMGSNASVTPDPVYVRMFPHGHGKA P DTSITFSVKVAYENGVEKELG ERFKVLPYKNCRNRTTKDFKFNKNKTAHIITVDPKTFEFVDGKLCITSRKPASEVPLAQRIY FF"	
ORIGIN		
Query Match	74.9%;	Score 2158.6; DB 2; Length 2407;
Best Local Similarity	93.6%;	Pred. No. 0;
Matches 2251; Conservative	0;	Mismatches 155; Indels 0; Gaps 0;
Qy	206	GTGAAACTCACACC CAAGAGCAGAAAAGTTCTTTGTATTATATTATCGCGCGCAAGTCGCT 265
Db	2	GTGAAACTCACACC CAAGAGCAAAAAGTTCTTTGTATTATATTATCGCGCGCAAGTCGCT 61
Qy	266	AGAAAGCGCAAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGT 325
Db	62	AGAAAGCGCAAAGCGAGGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGT 121
Qy	326	GCCCATATTATGAGCAGAACGCCGCCTGTGGAAAAAAAACCGTTGCCAGCTTATGCGAAG 385
Db	122	GCCCATATTATGAGCAGAGCGCGCGTGGCAAAAAGACCGTTGCGAACTTATGCGAAG 181

386 TGCATGCACCTTTTTCAGAAAAAGATCAAGTAATGCGGGGTGGGTAAATATATGTTCCCGAT 445
182 TGTATGCACCTTTTTCAGAAAAAGACAGGTATGTCGGGTGTGGGAAATATGTTCCCTGAT 241
446 CTAGGTGTAGAAAGCACCTTTCTCTGATGGTACGAAACTTGTAACTGTGAAATTTGGGCCCATC 505
242 TTAGCGGTGGAAGTACTTTTCCCGATGGCACCAACTCTGTTAAACCGTGAATTTGGCCCATC 301
506 GAACCCAGATGAGCACTTCAAAAGCGGGCGAAGTGAATAATTTGGTTGGATTAAGACATCGAG 565
302 GAACCCGATGAACACTTCAAAAGCGGGCGAAGTCAAAATTTGGCTGTGTAATAAGACATTTGAA 361
566 CTCAATGAGGCAAGAGTAACCGAACTTCAGGTACTAATTAAGAGGCCCTAAATCTCTTG 625
362 CTCAACCGAGTAAAGGAATTTACCGAATAGAGTTTACCAACGAAGGACCTTAAATCTCTTG 421
626 CATGTGGGTAGCCATTTCCACCTTTCTTGAAGCTTAAACGAAGGCACTTAAATTCGATCGTGA 685
422 CATGTGGGTAGCCATTTCCACCTTTCTTGAAGCCCAACGAAGGCACTTGAATTTGATCGGGAA 481
686 AAAGCCTATGSCAAACGCTTAGATATTTCCCTCTGGCAACACGCTACGCAATTTGGGCGAGGA 745
482 AAAGCCTATGSCAAACGCTTAGATATTTCCCTCTGGCAACACGCTACGCAATTTGGGCGAGGA 541
746 CAACCCGCAAGTGCAGTTGATTTCTCTTGGTGGCAGTAAATAAGTGAATTTGGGCAATGAAC 805
542 CAACCCGCTAAAGTGCAGTTAATCCCTCTTGGCGGCAGTAAATAAGTGAATTTGGGCAATGAAC 601
806 GGGCTTGTGAATACATCGCGGATGAACGCATTAACATTAAGCGCTTGAACAGGCGGAA 865
602 GGGCTTGTGAATATTTGCAAGTGAACGCATTAACATTAAGCGCTTGAACAGGCGGAA 661
866 TCTCACGGATTTTCAAGTAAAGAGACTTCCCATGAAATGAAAAACAAAGATATGTAAA 925
662 TCTCACGGATTTTCAAGTAAAGAGACTTCCCATGAAATGAAAAACAAAGATATGTAAA 721
926 TACCTACGACCCACCAAGCGGATTAAGTGCCTTTAGGAGATACCGATCTTTGGGCGAGA 985
722 TACCTACGACCTTACCACAGGCGCAAAAGTGCCTTTAGGAGATACCGATCTTTGGGCGAGA 781
986 AGTAGACATGACTATACCACTATGCGGAAGAACTTAAATTTGGCGGGGTAAACTAT 1045
782 AGTAGAACATGACTATACCACTTATGCGGAAGAGCTCAAAATTTGGCGGGGTAAACTAT 841
1046 CCGTAGGGGTATGGGTACAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTTCATCAC 1105
842 CCGTAGGGGATGGGTACAGAGCAATAGTCCAGATGAAAAACACCCCTAGATTTAGTTCATCAC 901
1106 TAAACGCGATGATTATCGACTACACCGGGATTTACAAAGCCGACATTTGGGATTTAAAAACGG 1165
902 CAACGCGATGATTATTTGACTTACACCGGGATTTACAAAGCCGACATTTGGCATTTAAAAATGG 961
1166 CAATAATCATGGCATTTGGCAAGGAGGAAACAGGACATGCAAGATGGGTAAAGCCCTCA 1225
962 CAATAATCATGGCATTTGGCAAGGAGGAGGAAACAGGACATGCAAGATGGGTAAAGCCCTCA 1021
1226 TATGTGTGGGTGGGTGGGCAAGAGCACTAGCAGGGGAAGGTATGATTTATCCGCTGG 1285
1022 TATGTGTGGGTGGGTGGGCAAGAGCACTTAGCAGGGGAAGGTATGATTTATCCGCTGG 1081
1286 GGGATTCGATTTCACACCCCACTTCTCTTCTTCCACAACTATCCCTACCGCTCTAGGCA 1345
1082 GGGATTCGATTTCACACCCCACTTCTCTTCTTCCACAACTATCCCTACCGCTCTAGGCA 1141
1346 TGGGTTTCAACCATGTTTGGAGGCGGCAAGGTCTCTGTAGATGGCAAGATGCGACTAC 1405
1142 TGGGTTTCAACCATGTTTGGCGGTGGCACAAGTCCCGGTAGATGGCAGCAATGCGACTAC 1201
1406 TATCACTCCGGGCAATGGAATTTGCACCGCATGTTTGGCGGCGACAGAGATTTCTAT 1465
1202 CATCACTCCGGGCAATGGAATTTGCACCGCATGTTTGGCGGCGACAGAGATTTCTAT 1261
1466 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACTTGTAGAAACAAGT 1525

1262 GAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCCAGTAAAAACAACTTGTAGAACAAAT 1321
1526 AGAAGCGGGCGGATTTGGTTTAAATTTGCATGAAGACTGGGGCACCAACCAAGTTCGAT 1585
1322 AGAAGCGGGCGGATCGGCTTTAAATTTGCATGAAGACTGGGGCACCAACTCCAAGTGCAT 1381
1586 CGATCACTGCTTTGAGGTGGCAGATGAATACGATGTGCAAGTTTGTATCACAACCGATAC 1645
1382 CGATCACTGCTTTGAGCGTAGCAGATGAATACGATGTGCAAGTTTGTATCACAACCGATAC 1441
1646 AGTCAATGAGGCAAGTTATGTAGATGACACCTTAATGCAATGAACGGGCGGCGCATCCA 1705
1442 GGTCAATGAGGCAAGTTATGTAGATGACACCTTGAATGCGATGAACGGGCGGCGCATCCA 1501
1706 TGCCTACCACTTTGAGGAGCGGTGGAGGACACTCACCTGATGTTTATCACCATTTGGCAGG 1765
1502 TGCTTACCACTTTGAGGAGCGGCGGAGGACACTCACCTGATGTTTATCACCATTTGGCAGG 1561
1766 CGAGCTCAATATTTCTAACCTCTCCACCCCGCTCTATCGCGGCTGAAGATGTGCTCCA 1825
1562 CGAGCTCAATATTTCTAACCTCTCCACCCCGCTCTATCGCGGCTGAAGATGTGCTCCA 1621
1826 TGCAGAACACTTTAGACATGCTCATGACATGCCACCACTTAGACAAACGATCCGCGAGGA 1885
1622 TGCAGAACACTTTAGACATGCTCATGACCTTGCACCACTTAGATAAAACGATCCGCGAGGA 1681
1886 TTTTCAATTTTCTCAAAAGCGGTATCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCA 1945
1682 TTTTCAATTTTCCAAAGCGGTATCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCA 1741
1946 TGATATGGGTGTGATCGCGATGACAAAGCTCGGATTTGCGAAGCAATGGGCGTGCAGCGA 2005
1742 TGATATTTGGCGGTATCGCGATGACAAAGCTCGGATTTGCGAAGCAATGGGCGGCTGCAG 1801
2006 AGTGATTTCTCGAACTTTGGCAGACTGCGGATGAAGATAAAGAAATTTTGGTGAAGCTTCC 2065
1802 AGTGATTTCTCGAACTTTGGCAGACTGCGGATGAAGATAAAGAAATTTTGGTGAAGCTTCC 1861
2066 TGAAGATGCGAAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATFACACTAT 2125
1862 TGAAGATGTCAGATAACGACACTTCCGTATCAACCGCTACATCTCCAAATFACACTAT 1921
2126 CAACCCCGCTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT 2185
1922 TAAACCCCGCTCTAACCCATGGGTAAAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAAAT 1981
2186 GCGCGACTTGTGTGTGAATCTTGCCTTTTGGCGGTAAAAACCCAAATCGTGATCAA 2245
1982 CGCTGATTTGGTGTGTGAATCTTGCCTTTTGGGTGTGAAACCTTAAGATTTGATCAA 2041
2246 AGGCGGTATGGGTGCTTCTCTGAAATGGGCGATTTCAACGCGTCTGTGCCCACTCCCA 2305
2042 AGGCGGTATGGGTGCTTCTCTGAAATGGGCGACTCCAAACGCGTCCGTGCCCTACACCTCA 2101
2306 ACCGGTTTATTTACCCCGAAATGTTTGGGCGCATCAACGCAAGGCGCAAAATTTGACACCAAGCAT 2365
2102 GCGGTTTATTTACCCCGAAATGTTTGGGCGCATCAACGCAAGGCGCAAAATTTGACACCAAGCAT 2161
2366 CACTTTTGTTCCTTCCAAAGTCCGCTATGAAAAATGGCGTGAAGAAAAAGCTGGGCTTAGAGCG 2425
2162 CACTTTTGTTCCTTCCAAAGTCCGCTATGAAAAATGGCGTGAAGAAAAAAGCTAGGCTTAGAGCG 2221
2426 CCAAGTTCTACGGGTCAAAAACTGCGGTAAACATCAACCAAGAAAGACTTCAAGTTCAACGA 2485
2222 CAAGGTGCTACCGGTGAAAAAATGCGCGCAACATCACTAAGAAAGACTTCAAAATTTCAACAA 2281
2486 CAAAACGCGCAAAAAATCACCGCTGATCCGAAACCTTTCGAGGTCTTTGTAGATGGCAAACT 2545
2282 CAAGAGGCGCATATCACTGTGATCTTAAACCTTTCGAGGTCTTTGTAGATGGCAAACT 2341
2546 CTGACCTCTTAAACCCACCTCGCAAGTGCCTCTTAGCCAGCGCTACACTTCTTCTTAGGC 2605

Db 2342 CTGCACCTCTAAACCCGGCTCTGAAGTGCCTCTAGCCAGCGCTACACTTCTCTTAGGC 2401
Qy 2606 ACAATG 2611
Db 2402 NCAATG 2407

RESULT 9
BD185304
LOCUS BD185304 2183 bp DNA linear PAT 17-JUN-2003
DEFINITION Helicobacter felis vaccine.
ACCESSION BD185304
VERSION BD185304.1 GI:31877504
KEYWORDS JP 2002355054-A/3.
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
1 (bases 1 to 2183)
Kusters, J.G. and Cattoli, G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: JP 2002355054-A 3 10-DEC-2002;
AKZO NOBEL NV
COMMENT OS Helicobacter felis
PN JP 2002355054-A/3
PD 10-DEC-2002
PF 16-JUL-2001 JP 2001214711
PR 17-JUL-2000 EP 0202565.8
PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI
PC
C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC
A61K39/23,
PC A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15,
PC C12N1/19,
PC
C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/ PC
53,
PC
G01N33/53, G01N33/566, G01N33/569// (C12N9/80, C12R1:01), (C12Q1/68, PC
C12R1:01),
PC C12N15/00, C12N5/00, A61K37/02
CC Helicobacter felis vaccine
FH Key Location/Qualifiers
FT CDS (3)..(683)
FT CDS (694)..(2181).
FEATURES
source
1..2183
/organism="Helicobacter felis"
/mol_type="genomic DNA"
/db_xref="taxon:214"

ORIGIN
Query Match 67.6%; Score 1948; DB 2; Length 2183;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 206 GTGAACCTCACACCCACAGAGCAAGAAAGTTCTTGTTATATATGCGGCGAAGTGGCT 265
Db 3 GTGAACCTCACACCCACAGAGCAAGAAAGTTCTTGTTATATATGCGGCGAAGTGGCT 62

Qy 266 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT 325
Db 63 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAATCAACCCGAGCCATTGCGCTACATTAGT 122

Qy 326 GCCCATATTAGCAGAAAGCGCGCTGGAAAAAACCCTTGCCCGAGCTTATGGAAGAG 385
Db 123 GCCCATATTAGCAGAGCGCGCTGGCAAAAAAACCGTTGCTGCAACTTATGGAAGAA 182

Qy 386 TGCATGCACTTTTGA AAAAGATCAAGTAAATGCCCGGGTGGTAAATATGTTCCCGAT 445
Db 183 TGTATGCACTTTTGA AAAAGATGAGGTATGCCCGGTGGGAAATATGTTCCCTGAT 242

Qy 446 CTAGGTGTAGAGCCACCTTTCTCTGATGCTACGAAACTTGTAACTGTGAATTTGGCCCATC 505

Db 243 TTGGCGGTAGAGCCACTTTCCCGATGGCACCAAACTCGTAACCGTGAATTGGCCCAATT 302
Qy 506 GAAACAGATGAGCACTTCAAAGCGGGCGAAGTAAATTTGGTTGCATAAAGACATCGAG 565
Db 303 GAACTCTGATGACACTTTAAAGCCGCTGAAGTAAATTTGGCTGTGATAAAGACATTGAG 362
Qy 566 CTCATAGCAGGCAAGAAAGTAAACCGAACTTGAAGTTACTAATGAAGGGCCTAAATCCTTG 625
Db 363 CTCACGTTGGGTAGAGAAAGTTACCGAGCTTGAAGTTACCAACGAAGGACCTAAATCCTTG 422
Qy 626 CATGTGGGTAGCCATTTCACCTTTCTTGAAGCTAACAGGCACTAAATTCGATCGTGA 685
Db 423 CATGTGGGTAGCCATTTCACCTTTCTTGAACCAACAGGCACTTGAATTCGATCGGGA 482
Qy 686 AAAGCTATGGCAACCGCTAGATATTCCTCTGGCAACACGCTACGCACTTGGGGCAGGA 745
Db 483 AAAGCTATGGCAACCGCTAGATATTCCTCTGGCAACACGCTACGCACTTGGGGCAGGA 542
Qy 746 CAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAGTGAATGGCATGAAC 805
Db 543 CAACCCGTTAAAGTGCAGTTAATCCCTCTTGGCGGTAGTAAAAAAGTGATTTGCATGAAC 602
Qy 806 GGGCTTGTGAATAACATCGCGATGAACGCCATAACATAAAGCGCTTGACAGGCGAAA 865
Db 603 GGGCTTGTGAATAATATTTCCGGAACGAAACGCCATAACATAAAGCACTAGACAAAGGCAAA 662
Qy 866 TCTCAGCGATTTATCAAGTAAAGGAGACTCCCATGAAATGAAAAAACAAGAAATATGTAAA 925
Db 663 TCTCAGCGATTCATCAAGTAAAGGAGACTCCCATGAAATGAAAAAACAAGAAATATGTAAA 722
Qy 926 TACCTACGGAACCCACAAAGCGATAAAGTGCCTTAGAGATACCGATCTTTTGGGCAGA 985
Db 723 CACCTACGGAACCCACCAAGCGATAAAGTGCCTTAGAGATACCGATCTTTTGGGCAGA 782
Qy 986 AGTAGAACAATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTAT 1045
Db 783 AGTAGAACAATGACTATACCACTATGGCGAAGAGCTCAAATTTGGCGCGGTAAAACTAT 842
Qy 1046 CCGTCAGGGTATGGGTGAGAGCAATAGCCCTCATGAAAAACCCCTAGATTAGTCAATCAC 1105
Db 843 CCGTCAGGGTATGGGTGAGAGCAATAGCCCAAGTAAACACCTTAGATTAGTCAATCAC 902
Qy 1106 TAAACGCGATGATTATCGACTACACCGGGATTTACAAAGCCGACATTTGGGATTAACACGG 1165
Db 903 CAAACGCGATGATTATCGACTACACCGGGATTTATAAAGCCGACATTTGGTATTAACAAATGG 962
Qy 1166 CAAATTCATGCGCATTTGCAAGGCAAGGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCA 1225
Db 963 CAAATTCATGCGCATTTGCAAGGCGGGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCA 1022
Qy 1226 TATGTCGTGGGTGGGCACAGAGCACTAGCAGGGGAAGGTATGATTATTAACCGCTGG 1285
Db 1023 TATGTCGTGGGTGGGCACAGAGCACTAGCAGGGGAAGGTATGATTATTAACCGCTGG 1082
Qy 1286 GGGAAATCGAATTCACACACCCACTTCCTTTCTCCCAACAAATTTCCCTACCGCTTAGCCAA 1345
Db 1083 GGGAAATCGAATTCACACACCCACTTCCTCTCTCCCAACAAATTTCCCTACCGCTTAGCCAA 1142
Qy 1346 TGGCGTTACAAACCATGTTTGGAGGCGGCACAGGTCCTGTAGATGGCAAGATCGGACTAC 1405
Db 1143 TGGCTGTACAAACCATGTTTGGAGGTGGCAAGGTCGGGTAGATGGCAAGATCGGACTAC 1202
Qy 1406 TATCACTCGGGCAATGGAACCTTCACCGCATGTTGGCGCAGCAGAGAGATTTCTAT 1465
Db 1203 CATCACTCGGGCAATGGAACCTTCACCGCATGTTGGCGCAGCAGAGATTTCTAT 1262
Qy 1466 GAATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAACAAAGT 1525
Db 1263 GAATGTAGGCTTTTGGGCAAGGCAATAGTTCTAGCAAAAAAACAATTTGTAGAACAAAGT 1322
Qy 1526 AGAAGCGGCGCATTTGGTFTTTAAATTTGATGAAGCTCGGGCAACCAACCAAGTGGCAT 1585

Db	1323	AGAAAGCGGCGGATTTGGCTTTAAATTCATGAAGACTGGGCGCAACACCAAGTGGCAT	1382
Qy	1586	CGATCACTGCTTGAGCGTGGCAGATGAATACATGTGCAAGTTGTATCCACACCGATAC	1645
Db	1383	CGATCACTGCTTGAGCGTGGCAGATGAATACATGTGCAAGTTGTATCCACACCGATAC	1442
Qy	1646	AGTCAATGAGCGAGTTATGTATGATGACACCTTAATGCAATGACGCGGCGCCATCCA	1705
Db	1443	GGTCAATGAGCGAGTTATGTATGATGACACCTTAATGCAATGACGCGGCGCCATCCA	1502
Qy	1706	TGCTTACACATTTGAGGAGCGGCTGGAGGACACTCACTGATGTATCACCATGGCAGG	1765
Db	1503	TGCTTACACATTTGAGGAGCGGCTGGAGGACACTCACTGATGTATCACCATGGCAGG	1562
Qy	1766	CGAGTCAATATTTACCTCTCTCCACACCCCACTATCCCTATACCATTAATACGGT	1825
Db	1563	CGAGTCAATATTTACCTCTCTCCACACCCCACTATCCCTATACCATTAATACGGT	1622
Qy	1826	TGCGAGACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCATCCGGGAGA	1885
Db	1623	TGCGAGACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCATCCGGGAGA	1682
Qy	1886	TTTACAAATTTCTCAAAGCCGTATCCGCCCGGCTCTATCCGGCTGAAGATGTCTCCA	1945
Db	1583	TTTACAAATTTCTCAAAGCCGTATCCGCCCGGCTCTATCCGGCTGAAGATGTCTCCA	1742
Qy	1946	TGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGGCTGCAGGGA	2005
Db	1743	TGATATGGGTGTGATCGCGATGACTAGCTCCGATTCCGCAAGCAATGGGGCTGCAGGGA	1802
Qy	2006	AGTGATTCCTCGAATCTGGCGAGACTGCGGATGAAGATAAAGAAATTTGGTAAGCTTCC	2065
Db	1803	AGTGATTCCTCGAATCTGGCGAATCTGCGAATGAAGATAAAGAAATTTGGTAAGCTTCC	1862
Qy	2066	TGAGATGGCAAGATTAACGATATTTCCGATTAAGCGCTACATCTCCAAATACACTAT	2125
Db	1863	TGAGATGGGTGAGATTAACGATATTTCCGATTAAGCGCTATATCTCCAAATACACTAT	1922
Qy	2126	CAACCCCGCTTTGACCAACCGCGTGGAGGATATATCGGCTCTGTGAAGAGGCGAAGAT	2185
Db	1923	TNATCCCGCTTTGACCAATCGCGTGGAGGATATATCGGCTCTGTGAAGAGGCGAAGAT	1982
Qy	2186	CGCGCATGCGTGTGTGAATCTCGCTTTTGGCGTAAACCCCAAAATCGTGATCAA	2245
Db	1983	CGCGCATGCGTGTGTGAATCTCGCTTTTGGCGTGAACCTTAAGATTGTGATTAA	2042
Qy	2246	AGCGGATATGCGTGTCTCTGAAATGGGCGATTTAAACGCTCTGTGCCCATCTCCCA	2305
Db	2043	AGTGCGCATGCGTGTCTCTGAAATGGGCGATTTAAACGCTCTGTGCCCATCTCCCA	2102
Qy	2306	ACCGGTTTATTACCGGAAATGTTTGGGCATCACGCAAGCGGCAAAATTTGACACCGCAT	2365
Db	2103	GCCGGTTTATTACCGGAAATGTTTGGGCATCACGCAAGCGGCAAAATTTGACACCGCAT	2162
Qy	2366	CACITTTTGTTCCTCAAGTCC 2385	
Db	2163	CACITTTTGTTCCTCAAGCG 2182	

RESULT 10
 AX356689
 LOCUS AX356689
 DEFINITION Sequence 7 from Patent EP1176192.
 ACCESSION AX356689
 VERSION AX356689.1 GI:18674026
 KEYWORDS
 SOURCE
 ORGANISM
 Helicobacter felis
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 Helicobacteraceae; Helicobacter.
 1
 Kusters,J.G. and Cattoli,G.
 Helicobacter felis vaccine
 TITLE

JOURNAL	Patent: EP 1176192-A 7 30-JAN-2002;
FEATURES	Akzo Nobel N.V. (NL)
source	Location/Qualifiers
	1..2183
	/organism="Helicobacter felis"
	/mol_type="unassigned DNA"
	/db_xref="taxon:214"
CDS	3..683
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD23259.1"
	/db_xref="GI:18674027"
	/translation="MKLTPKOEKFLYYAGBVARKRKAEGLKLNOPEAIAYISAHIM DEARKKTAELMEEFKFLKDDVPMGVPMVDPDLGVEATPDGDTKLVTVMPTPEP DEHFKAGVFKGCKDIIELNVGKEVTELEVTNEGPKSLHVGSHFHFETNKALKFPRE KAYGRLDI PSGNLTRIGAGTRKVLII PLGGSKKVI GNMGLVNNIADERHKHALDK AKSHGFIK"
	694..>2181
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD23260.1"
	/db_xref="GI:18674028"
	/translation="MKMKQEVNTYPTTGDKVRLGDTDLMAEVEHDTTYGSELKF GAGKTIREGMGQSNPDETLDLVITNAMI IDTYGIYKADI GIKNGKIHGIGKAGND MQDGVSFHMVVGVTGTEALAGEMI ITAGIDSHTHFLSPQFPPTALANGVTVMFGGT LPDGTNATITTPGKWLHRLMLRAABEYSMNQVFLGKNSKKOLVEOVEAGAI GFK LHEQDTPSALDHCLSVADYDVOVCIHTDTVNEAGVYDDTLNAMNGRAIHAYHILBG AGGSHSDVITMAGELNLPSTTPTI PTYINTVAEHLDMLTCHHLDRKIRREDLQFS AGRIKPSIAAEVDLHIGVIAMTSSDSQAMGRAGEVPRITWQTDADKNKKKFKLPED GADNDFIKRYI SYITINPALTHGVSEYIGSVEEGKIADLVNPNAPFPGVFKPIVIK GGMVVFSEMGDSNASVPTPQPVYVREMFGHGKAKPDTSIITPRVSS"
ORIGIN	
Query Match	67.6%; Score 1948; DB 2; Length 2183;
Best Local Similarity	93.3%; Pred. No. 0;
Matches 2035; Conservative	0; Mismatches 145; Indels 0; Gaps 0;
Qy	206 GTGAAATCTCACCCAAAGAGCAAGAAAGTTCTTGTATATATATGCGGCGCAAGTGGCT 265
Db	3 GTGAAATCTCACCCAAAGAGCAAGAAAGTTCTTGTATATATGCGGCGCAAGTGGCT 62
Qy	266 AGAAAGCCAAAGCAGAGGCTTAAAGCTCAACCAACCCGAGCCATTCCTTACATTAGT 325
Db	63 AGAAAGCCAAAGCAGAGGCTTAAAGCTCAATCAACCCGAGCCATTCCTTACATTAGT 122
Qy	326 GCCCATATTTAGACGAGCGCGCTGGAAAAAACCCTTGGCCCGCTTATGGAAGAG 385
Db	123 GCCCATATTTAGACGAGCGCGCTGGAAAAAACCCTTGGCTGCTCAACTTATGGAAGAA 182
Qy	386 TGCATGCACTTTTGAAGAAAGATGAAGTATGCGCGGGTGGTAAATATGTTCCCGAT 445
Db	183 TGTATGCACTTTTGAAGAAAGATGAGGTGATGCGCGGTGGGAAATATGTTCCCTGAT 242
Qy	446 CTAGTGTAGAACCCCTTTCTCATGTGTAGAAACTTGTAACTGTGAAATTTGGCCCATC 505
Db	243 TTGGCGGTAGAACCCCTTTCCCGATGGCACCACCTCGTAACCGTGAATTTGGCCCAT 302
Qy	506 GAACCATGATGACACTTCAAGCGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAG 565
Db	303 GAACCTGATGAACACTTTTAAAGCGGCTGAAGTGAATTTGGCTGTGTAAAGACATCGAG 362
Qy	566 CTCAATGAGCAAGAGTAAACCGAACTTGAGGTACTTAATGAAGGGCTTAATCTTCTG 625
Db	363 CTCAACGTGGGTAGGAAAGTTACCGAGCTTGAAGTTACCAACGAGGACCTTAATCTTCTG 422
Qy	626 CATGTGGTACCAATTTCCACTTCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 685
Db	423 CATGTGGTACCAATTTCCACTTCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 482
Qy	686 AAAGCCTATGGCAACCGCTAGATATTCCTCTGTGGCAACACGCTACGCAATGGGGCAGGA 745

Db 483 AAAGCCTATGGCAACGCCCTAGATATTCCTCTCTGGCAACAGCTACGCAITGGGGCAGGA 542
Qy 746 CAAACCCGCAAGTGCAGTGTGATTCCTCTTGGTGGCAGTAAAGAAAGTGAATGGCATGAAC 805
Db 543 CAAACCCGTAAGTGCAGTGTGATTCCTCTTGGCGGTAGTAAAGAAAGTGAATGGCATGAAC 602
Qy 806 GGGCTTGTGATTAACATCGCGGATGAAGCCATAAACATAAAGCGCTTGACAAAGCGGAA 865
Db 603 GGGCTTGTGATTAATATTTGGCGACGACGCGCATAAACAAAGCACTAGACAAGGCAAAA 662
Qy 866 TCTCAGCGATTTATCAAGTAAAGAGACTCCCATGAAATGAAAAAACAAGAAATATGTAAA 925
Db 663 TCTCAGCGATTCATAGTAAAGAGACTCCCATGAAATGAAAAAACAAGAGTATGTAAA 722
Qy 926 TACCTACGCGACCCACCAAGCGGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCAGA 985
Db 723 CACCTACGCGACCCACCAAGCGGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCAGA 782
Qy 986 AGTAGAATGACTATATACCATATGGGAAGAACTTTAAATTTGGCGGGGTAAAACTAT 1045
Db 783 AGTAGAATGACTATATACCATATGGCGAAGAGCTCAAAATTTGGCGGGGTAAAACTAT 842
Qy 1046 CCGTAGGGGTATGGTCAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCATCAC 1105
Db 843 CCGTAGGGGTATGGTCAGAGCAATAGCCCAAGATGAAAAACACCTTAGATTTAGTATCAC 902
Qy 1106 TAACGCGATGATATGCACTACACCGGATTTACAAAGCCGACATTTGGGATTTAAAAACGG 1165
Db 903 CAACGCGATGATATGCACTACACCGGATTTATAAAGCCGACATTTGGTATTTAAAAATGG 962
Qy 1166 CAAATCATGGCATTTGGCAAGGAGGAAACAAGGACATGCAAGATGCGTAAGCCCTCA 1225
Db 963 CAAATCATGGTATTTGGCAAGGCGGGGAAACAAGACATGCAAGATGCGTAAGCCCTCA 1022
Qy 1226 TATGGTGTGGGTGGGCACAGAGCACTAGCAGGGAAGGTATGATTTATTCGCTGG 1285
Db 1023 TATGGTGTGGGTGGGCACAGAGCACTAGCAGGGAAGGTATGATTTATTCGCTGG 1082
Qy 1286 GGGATCGATTTCACACACCCACTTCCTTTCTCCACAACAAATTCCTACCGCTAGCCAA 1345
Db 1083 GGGATCGATTTCGACACCCACTTCCTTCCTCCCAACAAATTCCTACCGCTAGCCAA 1142
Qy 1346 TGGCGTTACAAACCATTTTGGAGCGGCGACAGGTCCTGTAGATGGCAGAAATGGCACTAC 1405
Db 1143 TGGTGTACAAACCATTTTGGAGGTGGCAGAGTCCGCTAGATGGCAGAAATGGCACTAC 1202
Qy 1406 TATCACTCCGGGCAAAATGGAACTTGACCGCATGTTGGCGCAGCAGAGATTTCTAT 1465
Db 1203 CATCACTCCGGGCAAAATGGAACTTGACCGCATGTTGGCGCAGCTGAAGAGTATTTCTAT 1262
Qy 1466 GAATGTGGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAAAT 1525
Db 1263 GAATGTAGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAAAT 1322
Qy 1526 AGAAGCGGCGCGATTTGTTTTAAATTTGATGAAGACTTGGGCGCAACAACCAAGTGGCAT 1585
Db 1323 AGAAGCGGCGCGATTTGTTTTAAATTTGATGAAGACTTGGGCGCAACAACCAAGTGGCAT 1382
Qy 1586 CGATCACTGCTTGAAGCGTGGCAGATGAATAGATGTGCAAGTTGTATTCACACCGCATAC 1645
Db 1383 CGATCACTGCTTGAAGCGTGGCAGATGAATAGATGTGCAAGTTGTATTCACACCGCATAC 1442
Qy 1646 AGTCAATGAGCGAGTATGTAGATGACACCCCTAAATCAATGAACGGCGCGCCATCCA 1705
Db 1443 GGTCAATGAGCGAGTATGTAGATGACACCCCTAAATCAATGAACGGCGCGCCATCCA 1502
Qy 1706 TGCCTACCAATTTGAGGAGCGGTTGGAGGACACTCACCTGATGTTATCACCATGGCAGG 1765
Db 1503 TGCCTACCAATTTGAGGAGCGGTTGGAGGACACTCACCTGATGTTATCACCATGGCAGG 1562
Qy 1766 CGAGCTCAATTTCTACCTCTCTCCACCAACCCCACTATTCCTCATACCATTAATACGGT 1825
Db 1563 CGAGCTCAATTTCTACCTCTCTCCACCAACCCCACTATTCCTCATACCATTAATACGGT 1622

Qy 1826 TGCAGAAACACTTAGACATGCTCATGACATGCAATGCCACCACTTAGACAAACGCAATCGCGAGGA 1885
Db 1623 TGCAGAAACACTTAGACATGCTCATGACCTGCCACCACTTAGATAAGCGCATCGCGAGGA 1682
Qy 1886 TTTTCAATTTTCTCAAGCGGTATCCGCCCGCTCTATCGCGGTGAAGATGTGCTCCA 1945
Db 1683 TTTTCAATTTTCTCAAGCGGTATCCGCCCGGTATCTATTCGCGTGAAGATGTGCTCCA 1742
Qy 1946 TGATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGGA 2005
Db 1743 TGATATGGCGTGTGATCGCGATGACCTAGCTCCGATTCGCAAGCAATGGGGCGCTGGGA 1802
Qy 2006 AGTGATTCCTCGAACTTGGCAGACTGCGGATGAAGATAAAAAAGAAATTTGGTAAGCTTCC 2065
Db 1803 AGTGATTCCTAGAACTTGCAGAACTGCAGATAAGAAATAAAAAAGAAATTTGGTAAGCTTCC 1862
Qy 2066 TGAAGATGGCAAGATAAGATAATTTCCGGATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1863 TGAAGATGGTGCAGATAAGCAACTTCCGATCAACGCTATATCTCCAAATACACTAT 1922
Qy 2126 CAACCCGCTTTGACCCACGCGTGAAGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT 2185
Db 1923 TAATCCCGCTTTGACCCATGGCGTGAAGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT 1982
Qy 2186 CGCGCATGTTGGTGTGGAATCTGCTTTTGGCGTAAACCCAAATCGTGATCAA 2245
Db 1983 CGCGCATGTTGGTGTGGAATCTGCTTTTGGCGTGAACCTTAAGATTGTGATTA 2042
Qy 2246 AGGCGTATGTTGGTCTTCTGAAATGGCGATTTCAACGGCTGTGTCGCACTCCCCA 2305
Db 2043 AGTGGCATGTTGGTCTTCTGAAATGGCGATTTCAACGGCTCGTCCGCCCTCA 2102
Qy 2306 ACCGTTTATTACCCGAAATGTTTGGGCATCAACGCAAGCGAAATTTGACACCAAGCAT 2365
Db 2103 GCGGTTTATTACCCGAAATGTTTGGGCATCAACGCAAGCGAAATTTGACACCAAGCAT 2162
Qy 2366 CACTTTTGTTCCTCAAGTCG 2385
Db 2163 CACTTTTGTGTCCTCAAGCG 2182

RESULT 11
AF330621
LOCUS
DEFINITION
AF330621 8406 bp DNA linear BCT 13-MAY-2003
Helicobacter bizzozeronii tRNA ribosyltransferase gene, partial
cds; GidB, complete cds; urease gene cluster, complete sequence and
unknown gene.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AF330621.1 GI:27462193
Helicobacter bizzozeronii
Helicobacter bizzozeronii
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE
AUTHORS
Zhu, J., Teng, C.H., Chang, C.F., Chang, C.D., Simpson, K.W., Wei, C.,
McDonough, P., McDonough, S. and Chang, Y.F.
1 (bases 1 to 8406)
Cloning and characterization of a Helicobacter bizzozeronii urease
gene cluster
JOURNAL
DNA Seq. 13 (6), 321-331 (2002)
PUBMED
12652903
REFERENCE
AUTHORS
Zhu, J. and Chang, Y.F.
2 (bases 1 to 8406)
Direct Submission
JOURNAL
Submitted (21-DEC-2000) Population Medicine and Diagnostic Science,
College of Veterinary Medicine, Cornell University, Tower Road,
Ithaca, NY 14853, USA
FEATURES
Location/Qualifiers
1. 8406
/organism="Helicobacter bizzozeronii"
/mol_type="genomic DNA"
/db_xref="taxon:56877"

```

CDS
<1..929
/codon_start=3
/transl_table=11
/product="tRNA ribosyltransferase"
/protein_id="AAO15370.1"
/db_xref="GI:27462194"
/translation="LMVVERASNKLTHSTPEKIFDFPPPHALVVLNDDTKVILKARLPAH
TAQKRPVEILLHATSTACLAQNRKVRGLVLGLEGGYSCVEVLNGLRVLSPK
FGSGLLEWAGVLEMLHGVLPPLPKRDEAQLDQYQSVFAYKIPGAIAPATSLH
FSGSKHLLKRFVITLHVAGTFLSVQSEDIRHPHAEFCRLPQAGLTDEA
PYILCVGTALRAHYKRGFGKPCBLFLHLGNPKVHAQALLTNFHLPESTLMLVA
SMVLGDLKELYKLAIEBHYRFYSYGDGMLIL"
926..1465
/codon_start=1
/transl_table=11
/product="GidB"
/protein_id="AAO15371.1"
/db_xref="GI:27462195"
/translation="MRAKLDLFTLLLEWGSVHNLGAKNHQDIEHNIQDSLOVDFI
APFKSCDIGSGAGFPAPISLACSNARFILLEPNNAKVAFLHHVKLALNNLEIQR
MRIHVSPQSVLVDLITSRALMNAQNLIASLPREQGHFLFYKGSHLRTEIACADH
ECHYGRKRVFYQSQRDFA"
1670..2053
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAO15372.1"
/db_xref="GI:27462196"
/translation="MLVFGHPQIPCMPPKCVADIQAIQVPSHTIPFFQCTHTDAFGL
SKHCSQGVYAVVQNPLESCALMANKPFTYILAENQLEACQSIANDYLFDPARILKVI
EHAQLEEVUKRLDGVAFKSLWL"
2341..3045
/codon_start=1
/transl_table=11
/product="urease A"
/protein_id="AAO15373.1"
/db_xref="GI:27462197"
/translation="MKLTPEKLDKMLHYAGELAKRKANGVKLNTYEAVALLSAHVM
EEARAGKSVADLMQKRLTKADDPVMPGVAHMIHEVGIEANFPDGTGLVTIHTPVED
GKHLAPGEVLKNEIDITLNAQOATLVEVNHKGRDPVOVGSHPFPEVNKLLEFDR
KAYGKRLDIASCTAVRPEPEKKTVELIIGGNRIYGFNSLVDRQADTDGKKLALR
AKHGFGVNVCGDKK"
3061..4770
/codon_start=1
/transl_table=11
/product="urease B"
/protein_id="AAO15374.1"
/db_xref="GI:27462198"
/translation="MKKISKEYVSMGYPTTGDKVRLGDDTLILEVHEDCTTYGEEIK
FGGKTRIDGMAQTNPSHSELDVLNTNALIVDTYGIYKADIGIKNGKIHGIGKAGNK
DMODGVNNLCVGPATEALAAEGLIVTAGGIDTHIFISPOQIPTAPASGITWIGGG
TGPADGTNATTITPGSNLTKMLRASEEYAMNLGYLGKGNVSVEPSLVDOLBAGIIF
KIHEDNGSTPAALYHCLNLVADKVDQVAIHTDILNEAGCVEDTLQAIAGRTIHTFE
GAGGHAPDVRKWSGFNLPASTNPITPTVNTAEHMDMLVCHLDRKNIKEDVOF
ADSRIPRTAAEDKLHMDGIFPISITSDSQAMRGVEVITRTWTADKDKKBFGRLE
EKGDNDNFKRIYISKYITNPAHAGISYEVGSVEGKPADLWSPAPFGIKPNMI
KGGFIALSOMGDANASIPTPQPVYVYREMFHGHGKAKFDNIITPVSOVAYDNGIKERIG
LQSVLVPVKNCRNITKKDLKENDVTAHIEVNPETYKVKVDGKVTSCAADKISLAQLY
NLF"
4965..5552
/codon_start=1
/transl_table=11
/product="urease I"
/protein_id="AAO15375.1"
/db_xref="GI:27462199"
/translation="MLGLVLYVAIVLISNGVCLANVDARSKVMNVFVGGSIICN
VIAIVSTFNPTVCAEDVAQVQSOLHIFNYGPATGLLFGFTYMAAIIINNIYLDWK
PYGWYCLFVINTIPAILSHYDALDHRLLGITEGDWWAFITLWANGVWLWLTGWBIC
ALGKSLGKVPWLAIIEGVTITAIPAWLLFIQHS"
5563..6093
/codon_start=1
/transl_table=11
/product="urease E"

```

```

/protein_id="AAO15376.1"
/db_xref="GI:27462200"
/translation="MLASSVLGNLKBGSSKELDFIDLEWFDQAKRMGRFTSQKGAEI
VLKKNPPMKCDGDIILFEDATSLIAINIIPTPLHYVADSTQVARIICYEVGNRHA
SLYGDSPLSFKTPPRLQVLFDKLALRYEVLKSKLDASQRISSVAPHADPLQEGSA
PLKFKSAULDLQIVIKK"
6098..6775
/codon_start=1
/transl_table=11
/product="urease F"
/protein_id="AAO15377.1"
/db_xref="GI:27462201"
/translation="MQKDFLLQVNDAMPFISYTHSGLETYIOHKEYSSKESALDY
MRAYSTOFLYTELLALXKLAGYAHQEDLESLEMEEQICLATPPLELSANQKLGNR
FLKTIGVLDPFLSAFFKDYMQKSTPTTHASAYFCACLGLEASLKHLYLAQSNM
VINCVKTIPLAQNDQRIILLALQETFTTLLVTLEGLDASYLCAASIQNDIKAMQHEHL
YSRLYMS"
6797..7396
/codon_start=1
/transl_table=11
/product="urease G"
/protein_id="AAO15378.1"
/db_xref="GI:27462202"
/translation="MWKIGVCGPVSGSKTALIEALTRAMSOQYSIGVITNDIYTKEDA
EFLCRNSVMPRERIGVETGCPHTAIREDSAMNLEAVEELHAKFPDIELMFIESGD
NLSTTFNPELADFTIFVIDVAEGDKIPRKGGGPIRSDLLIINKIDLAPYVGADLGM
DRDSKQKRGDKPFLFTNIRSKGLDHVIGWIKKYALLED"
7399..8232
/codon_start=1
/transl_table=11
/product="urease H"
/protein_id="AAO15379.1"
/db_xref="GI:27462203"
/translation="MAVOHPTSYAQAPRLYLRKTIQONGKVCVIADNYSPPEKLMPPF
YERGNMAEILIIAVSPGMLKGAQOIQMDIGENCOLELSFOSFKIQDTEDGVATENT
QIHHPNALLDFSPPLPIPPANRPFKSHSHVLHKNQALLYSEIITAGRIANGEPAF
HKIHSTIKISTQEGESLSRVLDNTILEPASMDLKNPCWFGNYTHLIFLRLDKL
EDLLELLESTQINAGVSLPRLPEGYHSLCLKALANGSEPLLALRKISINLLVERME
G"

```

ORIGIN

Qy	Db	Query Match	Best Local Similarity	Matches 1723;	Conservative	Score 1135.6;	DB 15;	Length 8406;
73	TATTATTAAATTTTAAATAATTTACTTATATCATATAATAATAATTTACTTATATTA 132	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
2208	TAAGATCAACTAGATAATAATAATTAATCTTGTCTAGAAAATAATTAACAAGGTAGC 2267	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
133	AAAAAGTTAAATAAAAGTAACGAAATTAGGACTATAATCCCATTCCTTTTAAATAATTAA 192	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
2268	CAAAACAGACTAGAAATTTGTCCCGTTGATAGCTAGCTATTCAATAATACAAATTTGGTA 2327	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
193	CAAGAGTAATAGGTGAACTCACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 252	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
2328	GAAGGAGTTTAGGATGAAATTAACCCCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2387	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
253	GGGCGAAGTGGCTAGAACCGCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
2388	GGGCGAATTTGGCTAAAAAAGCAAGCAAAATGGCGTTAAAGCTTAAATTAATTAAGTGG 2447	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
313	TGCTTTACATTAAGTGGCCCATATTATGGACGAAAGCGCGCGTGGAAAAAACCCTTTGCC 372	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
2448	AGCCTCATAGTGGCCCATGTGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2507	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
373	GCTTATGGAAGAGTGCATGCATCTTTTGAAGAAAAAGATGAAGTAATGCCCCGGGTGGTAA 432	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
2508	TTTGATCAAG 2567	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
433	TATGGTTCCTGATCTAGGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492	39.4%	67.2%	0;	0;	1135.6;	15;	8406;
2568	TATGATCCACGAAGTGGGGATTGAAGCTTAACCTTCCTGATGGGACAAACCTGGTAACCAT 2627	39.4%	67.2%	0;	0;	1135.6;	15;	8406;

QY 493 GAATTGGCCCATCGAACCAGATGAGCACTTCAAAGCGGCGAAGTGAATTTGGTTGCGA 552
Db 2628 CCATACCCCGCTGAAGATGTGGGCAATAAATGGCTCCGGGTGAAGTGAATTTGAAAAA 2687
QY 553 TAAAGACATCGAGCTCAATGCAAGCAAGAAAGTAAACCGAATCTGAGGTTTACTTAATGAAGG 612
Db 2688 CGAAGACATCACTTTGAATGCAAGCAAGAAAGTAAACCGAATCTGAGGTTTACTTAATGAAGG 2747
QY 613 GCCTAAATCTTGTGATGGGTAGCCATTTCCACTTCTTTGAAGCTAAACAGGCACTAAA 672
Db 2748 CGATCGCCCGCTGCAAGTGGGCTCCCACTTCCACTTCTTTGAAGTGAATAAGCTTTTGGG 2807
QY 673 ATTGCATCGTGAAGAGCTATGCAAGAGCGCTAGATATTCCCTCTGCAACACGCTAGG 732
Db 2808 ATTTGATCGTGAAGAGCTATGCAAGAGCGCTAGATATTCCCTCTGCAAGAGCGCTAGG 2867
QY 733 CATTTGGGCGAGCAAAACCCGCAAGTGCAGTTGATTCCTCTGCTGGCGAGTAAAAAGT 792
Db 2868 CTTTGAACCCGCTGGAAGAAAAACCGTGGAAATTTGAATTAAGTTGGCGTAAACAGCAT 2927
QY 793 GATTGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAAACATAAAGCGCT 852
Db 2928 TTACGGCTTTAACTCTCTTGTGGATCGCAAGCGGCTAGATATTCCCTCTGCAAGAGCGCTAGG 2987
QY 853 TGACAGCGGAAATCTCAGGATTT-----ATCAAG 883
Db 2988 CAAACGCGCCAAAGAACATGGCTTTGGTGTGTGAAATTTGGGCTTGGCATAAAAATAAGG 3047
QY 884 TAAGGAGACTCCCATGAAA---ATGAAAAACAAGATATGTAATATCTACGAGCCAC 940
Db 3048 AAAGGACAAATCCGATGAAAAAAATCTCGAAAAAGATATGTTCTATGATGGGTAACCAAGCAT 3107
QY 941 CAAAGCGCATAAAGTGCCTTAGGAGATACCGATCTTTGGGCGAGAGTAGAACATGACTA 1000
Db 3108 TACGGCGATAAGTGAATTTGGGCGATACCGACTGTCTTAGAAGTCGAACATGACTG 3167
QY 1001 TACCCTATGCGGAAGAACTTAAATTTGGCGGGGTAAACTATCCGTGAGGGTATGGG 1060
Db 3168 CACCACCTATGCGGAAGAAATTAAGTTGTGGCGGTAAAAACCAATCGCGATGGGATGGC 3227
QY 1061 TCAGAGCAATAGCCTGATGAACACACCTAGATTTAGTCACTACATACCGCATGATAT 1120
Db 3228 ACAACCAACAGCCCGACGACCACTGATCTTGTGCTCACTAACGCCCTGATCGT 3287
QY 1121 CGACTACACCGGATTTACAAAGCGCATTTGGGATTAACAAACGCAAAATCCATGGCAT 1180
Db 3288 GGAATTACACCGCATTTAAAGCGCATTTGGCAATTAACAAATCCATGGCAT 3347
QY 1181 TGGCAAGCGAGAAACAAGGACATGCAAGATGGCGTAAGCCCTCATATGGTCTGGGGTGT 1240
Db 3348 TGGCAAGCGAGCAATAAAGACATGCAAGATGGCGTTTGAACAATCTTTGGGTGGGCC 3407
QY 1241 GGGCACAAGACATGACAGGGAAGGATATGATTAATACCGCTGGGGATCGATTCACA 1300
Db 3408 TGCTACTAGGGCTTTGGCGCTGAAAGGGCTGATTTGACGTGTGGGATTTGACACCCA 3467
QY 1301 CACCACCTTCTTCTCCACAACAATTCCTACCGCTCTAGCCCAATGGGTTTACAACCAT 1360
Db 3468 CATCACCTTTATTTCTCCCAACAATCCCAAGCATTTGCGGGGATCACACCAT 3527
QY 1361 GTTTGGAGCGGCACAGGTCTCTGATGGCAAGATGCGACTACTATCATCTCCGGGCA 1420
Db 3528 GATTGGTGGGGAACAGGTCCAGCTGATGGGACTAACCGCATACCATCTCCTCCGGGGC 3587
QY 1421 ATGGAATTTGACCCGATTTGGCGGCGAGGAGAGTATTTCTATGAATTTGGGCTTTT 1480
Db 3588 CTGGAACCTTTAAACCATGCTCCGTGCTCTGAGAAATATGCAATTTTGGGCTATT 3647
QY 1481 GGGCAAGGCAATAGCTTAGCAAAACAACTCTGTAGAACAGTAGAAGCGGGCGCAT 1540
Db 3648 GGGTAAAGGGAAATGTCTTTATGAACCCCTCCCTGCTGATCAACTCGAAGCTGGAGCCAT 3707
QY 1541 TGGTTTTAAATGCAATGAGACTGGGGCAACACCAAGTGGGATCGATCACTGCTTGAG 1600

RESULT 12

Db 3708 TGGCTTTAAAAATCCAGAAAGCTGGGGTAGCACACCTGCAGCCATCTACATTTGCTTGA 3767
QY 1601 CQTGCGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGGATACAGTCAATCAGGCGAGG 1660
Db 3768 TGTGCTGCAATACGATGTGCAAGTGGCTATCCACACCGGATACCTTGAATGAGGGG 3827
QY 1661 TTATGTAGATGACACCTTAAATGCAATGAACGGGGCGGCATCCATGCTTACCAATGTA 1720
Db 3828 CTGTGTGGAAGACACTTTTGAAGCCATTTGCTGGGGCACTATCCACACTTTCCACACTGA 3887
QY 1721 GGAAGGGGTGAGGACACTCACCCTGATGTATCACCATGAGGCGAGGCTCAATTTCT 1780
Db 3888 AGGTGCTGTGTGGGGCAGCTCCGGATGTCAATGAATGTCTGGCGAATTTTAAACATCT 3947
QY 1781 ACCCTCTCTCACACACCCCACTATTCCCTATATACCAATTAATACGTTGCGAGAACACTTAGA 1840
Db 3948 CCCAGCTTCTACCAACCCCACTTCTTCTTACCGGTGATACAGAAGCCGAACACATGGA 4007
QY 1841 CATGCTCATGACATGCAACCTAGACAAACGCAATCCGCGAGGATTTACATTTTCTCA 1900
Db 4008 CATGTTGATGTGTGCCACCACTTGGATAAAAAACATCAAGAAGATGTCCAGTTTGTGTA 4067
QY 1901 AAGCCGTATCCGCGCGCTCTATCGCGCTGAAGATGTCTCCATGATATGGTGTGAT 1960
Db 4068 TTCTAGGATTTCCGCGCGCAACCATCGCGCTGAGGACAACTCCAGCATATGGGATTTT 4127
QY 1961 CCGCATGACAAGCTCGGATTTCCAAAGCAATGGGGGTGCGAGCGCAAGTATTCCTCGAAC 2020
Db 4128 CTCTATCACAGCTCTGACTCCCAAGCGATGGCGGTGAGCGAGGTCAACACCGCAC 4187
QY 2021 TTGGCAGACTGCGGATAAGATAAAAAAGAAATTTGGTAACTTCTGTAAGATGGCAAGA 2080
Db 4188 TTGGCAAAACAGCGGCAAAAAAAGAAATTTGGTGTCTGCTGAGGAAAAAGCGCA 4247
QY 2081 TAAACGATTAATTTCCGATTAAGCGGTAGATCTCCAAATACACTCAACCCCGCTTGAC 2140
Db 4248 TAATGACAACTTTCCGATCAAGCGCTACATTTCCAAATACACCATCAAGCCCGCTATTGC 4307
QY 2141 CCACGCGGTGAGCGATATATCGGCTCTGTGAAGAGGCAAGATCGCGCACTTTGGTGT 2200
Db 4308 ACACGCGATTTCTGAATATGTGCGCTCTGTAGAGTGGGCAAAATTCGCCGATTTGGTCT 4367
QY 2201 GTGGAATCTGCTTTTGGGCTTAAACCCAAATCGTATCAAAAGCGGATATGGTGT 2260
Db 4368 TTGGAGTCTGCTTCTTTGGCATTAACCCCAACATGATCATCAAAGCGGATTCATCGC 4427
QY 2261 CTTCTCTGAATGGGCGATTTAAACGCGTCTGTGCCACTCCCAACCGGTTTATTACCG 2320
Db 4428 ACTTTCTCAATGGGCGATGCCAATGCTCTATCCCACTCCCAACCGGTTATTACCG 4487
QY 2321 CGAAATGTTTGGGCACTACGGCAAGCGAAATTTTGACACAGCATCACTTTTGTTCCAA 2380
Db 4488 CGAAATGTTTGGCCACCTATGGTAAAGCCAAATTTGACCAATATCACATTTGTATCCCA 4547
QY 2381 AGTCGCTATGAAATGGGCTGAAAGAAAGCTGGGCTTTAGAGCGGCAAGTTCTACCGGT 2440
Db 4548 AGTGGCTTATGACACCGCATTTAAGAAGATTTGGCTTTGCAAGAGTGGTTTGGCAGT 4607
QY 2441 CAAAACTCGCGTAAACATCAACGAAGAACCTTCAAGTTCAACGACAAACGCGCAAAAT 2500
Db 4608 TAAAACTCGCGCAACATCAACAAAAAGACCTCAAAATCAACGATGTACCAGCACAT 4667
QY 2501 CACCGTGCATCCGAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCAACCTCTTAAAC 2560
Db 4668 CGAAGTCAATCTGAAACCTCAAAAGTTAAAGTGAAGTGGCAAGAGGTTACTTCCAAAGC 4727
QY 2561 CACCTCGCAAGTGCCTCTAGCCGAGCGCTACACTTTTCTTCTAGG 2604
Db 4728 AGCGGATAAAATCAGCCTAGCACAACTCTACAACTTTGTTCTAGG 4771

HECUREASE
LOCUS HECUREASE 2664 bp DNA linear BCT 03-MAY-1995
DEFINITION Helicobacter heilmannii urease, complete cds's.
ACCESSION L25079
VERSION L25079.1 GI:793908
KEYWORDS urease.
SOURCE Candidatus Helicobacter heilmannii
ORGANISM Candidatus Helicobacter heilmannii
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
1 (bases 1 to 2664)
Solnick,J.V., O'Rourke,J., Lee,A. and Tompkins,L.S.
Molecular analysis of urease genes from a newly identified
uncultured species of Helicobacter
Infect. Immun. 62 (5), 1631-1638 (1994)
JOURNAL
PUBMED 8168924
COMMENT
On May 3, 1995 this sequence version replaced gi:529422.
Original source text: Helicobacter heilmannii (individual isolate
2) DNA.

FEATURES
source
Location/Qualifiers
1..2664
/organism="Candidatus Helicobacter heilmannii"
/mol_type="genomic DNA"
/isolate="2"
/db_xref="taxon:35817"
/lab_host="BALB/c"
200..204
211..915
/codon_start=1
/transl_table=11
/product="urease"
/protein_id="AAA65722.1"
/db_xref="GI:529423"
/trãnslation="MKLTPEKLDKMLHYAGELAKQKAKGIKLYNTEAVALISAHV
EEARAKKSVADLMQGRILLKADDVMPGVAHMHEVGEIAGPDGTLVTHHPVEA
GSDKLAPVEVLKNEKLLINAGKHAVOLKVKNGDRPVQVGHSHFPEFVKNLKDLPDE
KATGKRLDASGTAVAFPEFGEETVELIDIGNKRIYGFNALVDRQADHGKKLALKR
AKEKHFTINGCGDNK"
919..923
RBS
930..2636
CODS
/codon_start=1
/transl_table=11
/product="urease"
/protein_id="AAA65723.1"
/db_xref="GI:529424"
/trãnslation="MKKISKEVYSVMGPTTGDKVRLGDTDLILEVEHDCTTYGEIK
FGGKTRIDMGMTNPSPSHLELDVITNALIVDTYGIYKADIGIRKNGKIHGIGKAGN
DLQDGVNRLCVGPTEAKLAAGLIVTAGGIDTHIFISPOQIPTAFASGITTMIGGG
TPGADTNATITPGRWNLKEMLRSEBYAMNLGYLGKGNVSPFPAIDQLBAGATG
KIHEDGSPVPSALINHALINDADKYDVQVAIHTDTLNBAAGCVDTLEAIAGRTTHTPHE
GAGGHAPDPSIKWAGFNILPASTNPTIPTKNTAEHMDMLMCHLDKNIKEDVEFA
DSRIRPQTAAEDKLDHMGIFSLTSSDSQAMGRVGEIVTRQTADKNKKEFGRLPEE
KGDNDNFRIKRYISKYITINPAITHGISEYSGVSEVGEKYADLVLSPPAFGIKENMLIK
GGFIALSQMGDANASITPPQVPVYREMFHGHGKAKFDNTITFVSQVAYENGIKHELJ
QRVLPVKNCRNITTKDKLFENDVTAHIEVNPETYKVKVDGNEVTSHAADKLSLAQLYN
LP"

ORIGIN

	Query Match	38.2%	Score 1099.8;	DB 15;	Length 2664;
	Best Local Similarity	67.0%;	Pred. No. 0;		
	Matches 1722;	Conservative	1;	Mismatches 813;	Indels 36; Gaps 9;
QY	63	TTGTTAATRCTATTATTTT	TTTAAATATTAATCTATTAT	TATCATATATATAATATATT	122
Db	70				
QY	123	ACTTATATTAAAAAGTTAAT	AAAAAGTAACGAAATTAGGACT	TATATCCCATTCGCCTTTTA	182
Db	129	ACCATTAGCCAAAGATAC	TAAATCTGCCTCGTGTAGCT	TGGCGCTATTCAATAACA	187
QY	183	AAATTTAAACACAGGAGT	TAATAGGTGAAACTCACACCC	AAAGACGAAGAAAAGTTCTTGT	242
Db	188	CCATTTTATACGAGGAGT	TTAGATGAAACTGACACTAA	GAGTTGGATAAGTTGATGC	247

1292 CGATTTCACACCCACTTCTCTTCTCCACAACAATTCCTTACCGCTCTAGCCATGGCGT 1351
1328 CGACACCCACATCCATCTTCTCTCCACAACAATTCCTTACCGCTCTAGCCATGGCGT 1387
1352 TACAACATGTTTGGAGCGGCACAGGTCCTGTAGATGGCAGCAATGGAGTACTATCAC 1411
1388 CACACCATGATCGGGGGGGGACTGACCTGACATGGGACCAACGCGCACCATCAC 1447
1412 TCCGGGCAATGGAACTTGCACCGCATGTTGCGGCGAGCAGAGAAGTATTTCTGAATGT 1471
1448 TCCGGCGCGCTGGAACTTAAAGAAATGCTCGCGCTTCTGAAGAAATACGCCATGAACCT 1507
1472 GGGCTTTTGGGCAAGCAATAGCTCTAGCAAAAAACAATTTAGAACAAAGTAGAAGC 1531
1508 TGGCTACCTTGGTAAAGGGGAATGTCTCTTTTGAACCTGCTCTCAATGACCAAGCTAGAAC 1567
1532 GGGCGCATGTTTAAATTTGAATGCAATGCAAGCTGGGGCAACAACCAAGTCGATGATCA 1591
1568 CGGCGCATGTTTAAATTTGAATGCAAGCTGGGGCAACAACCTCAGGCATCAACCA 1627
1592 CTGCTTGAAGCGTGCAGATGAATGATGTCGAAGTTTGTATCCACACCGCATACAGTCAA 1651
1628 CGCTCTAAACATCGCTGACAAATACGATGTCGAAGTGGCGATCCACACCGACCTTGAA 1687
1652 TGAGGCGAGTTATGTAGATGACACCTTAAATGCAATGAACGGGCGCGCATCCATGCCCTA 1711
1688 TGAAGCGCGTGTGTGAAGACACCTTGAAGCCATCGCTGGAGCGACCATCCACACCTT 1747
1712 CCACATTTGAGGCGGGTGGAGGACACTCACTGATGTTATCCATGAGCGAGCGT 1771
1748 CCACACGGAAGTGTGGCGGGGACACGCTTCTGACGTGATCAAAATGGGCGGCGAAT 1807
1772 CAATATTTCTACCTTCTCCACACCCCACTATTTCCCTATACCAATTAATACGTTGCGAGA 1831
1808 TAAATCTTCTGCTTCTACCAACCCCACTATTTCCCTTCCACCAAAACACAGAGCGGA 1867
1832 ACATTTAGATGTCATGATGACATGACCACTAGACAAACGCAATCCGCGAGGATTTACA 1891
1868 ACATGACATGCTTATG---TGCACCACTTGGATAAAAACATCAAAAGATGTGGA 1924
1892 ATTTTCTCAAGCGGTATCCGCGCGGCTCTATCGCGCTGAAGTGTCTCCATGAT 1951
1925 ATTTGCGGACTTCAGTATCCGCGCGGCAACCAATTTGCGCGGAAGATAAATCCACGACAT 1984
1952 GGGTGTGATCGGATGACAAGCTGGATTCGCAAGCAATGGGGTGCAGCGCAAGTCAAT 2011
1985 GGGATTTTCTCCTCCTACTACTCGACTCTCAAGCCATGGTCCGCTGGGCGAAGTCAAT 2044
2012 TCCTCGAACTTGGCAGACTGCGGATTAAGAAATTAAGAAATTTGGTAAAGCTTCCCTGAAGA 2071
2045 CACCGGCACTTGGCAACAGCGGACAAACCAAAAGAAATTTGGCGCTTGGCTGGAAGA 2104
2072 TGGCAAGATTAAGATTTTCGATTAAGGCTATGATCTCCAAATACACTATCAACCC 2131
2105 AAAAGGCGACACGACATTTCCGATCAACGCTATATTTTCCAAATACACCATCAACCC 2164
2132 CGCTTTGACCCACGCGTGAAGATATCGGCTCTGGAAGAGGCGCAAGATCGCGGA 2191
2165 TGCCATCACACGGCATTTCTGATATATGTCGGCTCTGTAGAGTGGGTAATACGCTGA 2224
2192 CTGCTGTGTGGAACTCTGCTCTTTTGGGCTTAAACCCAAATTCGTGATCAAGGCGG 2251
2225 CTGTGTCTTGGAGCCCTGCTCTTCTTGGCATTAACCCCAACATGATCATCAAGGCGG 2284
2252 TATGCTGTCTTCTGAAATGGGCGATTTCAACGCGCTGTGTGCCATCTCCCAACCGGT 2311
2285 TTTTATTTGCGCTTCTCAAAATGGGCGATGCGCAACGCTTCTATCCCACTCTCCCAACCGGT 2344
2312 TTATTTACCGCAATTTTGGCATCGGCAAGCGGAATTTGACACCGACATCACTTT 2371
2345 GTACTACCGCAATTTTGGCGGCAACCGGCAAGCGGAATTTGACACCAACATCACTTT 2404
2372 TGTTTCCAAAGTGCCTATGAAATGGGCTGAAAGAAAGTGGGCTTAGAGCGCGCAAGT 2431

Db 2405 CGTGTCTCAAGTGGCTTATGAAAACGGCATTAACACAGATTTGGGCTTCAAGAGATTGT 2464
Qy 2432 TCTACCGGTCAAAAACCTGCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAAAC 2491
Db 2465 GTTGCTGTGAAAACCTGCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGATGTCAC 2524
Qy 2492 GCAAAATACCGCTGCATCGAATCCGAAACCTTCGAGGCTTTTGTAGATGGCAAACTCTGCAC 2551
Db 2525 CGCACATCGAAGTCAACCTTGAACCTTACAAAGTGAAGTGGATGGCAACGAAGTTAC 2584
Qy 2552 CTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTTCTAG 2603
Db 2585 CTCCTCATCGGCTGACAAATTTAGCTTACACACTCTATTAACCTTATTTAG 2636

RESULT 13
AE001439_00/c
WPCOMMENT

Sequence split into 17 fragments LOCUS AE001439 Accession AE001439

Fragment Name	Begin	End
AE001439_00	1	110000
AE001439_01	100001	210000
AE001439_02	200001	310000
AE001439_03	300001	410000
AE001439_04	400001	510000
AE001439_05	500001	610000
AE001439_06	600001	710000
AE001439_07	700001	810000
AE001439_08	800001	910000
AE001439_09	900001	1010000
AE001439_10	1000001	1110000
AE001439_11	1100001	1210000
AE001439_12	1200001	1310000
AE001439_13	1300001	1410000
AE001439_14	1400001	1510000
AE001439_15	1500001	1610000
AE001439_16	1600001	1643831

LOCUS AE001439 1643831 bp DNA circular BCT 19-JAN-2006

DEFINITION Helicobacter pylori J99, complete genome.

ACCESSION AE001439 AE001440-AE001571

VERSION AE001439.1 GI:12057207

KEYWORDS

SOURCE Helicobacter pylori J99

ORGANISM Helicobacter pylori J99

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.

REFERENCE 1 (bases 1 to 1643831)

AUTHORS Alm, R.A., Ling, L.-S.L., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C., Smith, D.R., Noonan, B., Guild, B.C., deJonge, B.L., Carmel, G., Tummino, P.J., Caruso, A., Uria-Nickelsen, M., Mills, D.M., Ives, C., Gibson, R., Merberg, D., Mills, S.D., Jiang, Q., Taylor, D.E., Vovis, G.F. and Trust, T.J.

TITLE Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori

JOURNAL Nature 397 (6715), 176-180 (1999)

PUBMED 9923682

REFERENCES 2 (bases 1 to 1643831)

AUTHORS King, B.L., Alm, R.A. and Trust, T.J.

TITLE Direct Submission

JOURNAL Submitted (12-JAN-1999) Astra Research Center Boston, 128 Sidney Street, Cambridge, MA 02139, USA

COMMENT On or before Jan 19, 2006 this sequence version replaced gi:4154501, gi:4154511, gi:4154523, gi:4154537, gi:4154549, gi:4154559, gi:4154573, gi:4154583, gi:4154594, gi:4154604, gi:4154617, gi:4154626, gi:4154639, gi:4154651, gi:4154666, gi:4154678, gi:4154689, gi:4154703, gi:4154713, gi:4154723, gi:4154733, gi:4154749, gi:4154760, gi:4154775, gi:4154789, gi:4154799, gi:4154813, gi:4154826, gi:4154838, gi:4154861, gi:4154869, gi:4154880, gi:4154897, gi:4154910, gi:4154929, gi:4154939, gi:4154949, gi:4154961, gi:4154971, gi:4154979, gi:4154992, gi:4155005, gi:4155018, gi:4155034, gi:4155043, gi:4155057, gi:4155074, gi:4155086, gi:4155100, gi:4155102,

gi:4155115, gi:4155127, gi:4155147, gi:4155161, gi:4155170, gi:4155179, gi:4155191, gi:4155199, gi:4155211, gi:4155220, gi:4155238, gi:4155252, gi:4155264, gi:4155275, gi:4155285, gi:4155295, gi:4155310, gi:4155326, gi:4155338, gi:4155350, gi:4155371, gi:4155382, gi:4155392, gi:4155407, gi:4155418, gi:4155425, gi:4155431, gi:4155444, gi:4155454, gi:4155471, gi:4155483, gi:4155492, gi:4155505, gi:4155513, gi:4155523, gi:4155533, gi:4155550, gi:4155558, gi:4155575, gi:4155590, gi:4155602, gi:4155617, gi:4155628, gi:4155636, gi:4155656, gi:4155664, gi:4155675, gi:4155687, gi:4155697, gi:4155706, gi:4155717, gi:4155724, gi:4155739, gi:4155753, gi:4155763, gi:4155776, gi:4155799, gi:4155811, gi:4155845, gi:4155858, gi:4155872, gi:4155887, gi:4155898, gi:4155906, gi:4155914, gi:4155929, gi:4155938, gi:4155952, gi:4155964, gi:4155974, gi:4156044, gi:4156000, gi:4156017, gi:4156027, gi:4156032, gi:4156044, gi:4156051, gi:4156065, gi:4156083, gi:4156095, gi:4156108, gi:4156120.
 Addresses all correspondence to: hp@rcb.us.astro.com or Richard A. Alm, Astra Research Center Boston, 128 Sidney Street, Cambridge, MA, 02139. Lo-See L. Ling, Donald T. Moir, Douglas R. Smith, Braydon C. Guild, Gilles Carmel, Anthony Caruso, Debra M. Mills, Rene Gibson, and Gerald F. Vovis are with Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA, 02453. Qin Jiang and Diane E. Taylor are with the University of Alberta Department of Medical Microbiology and Immunology, Edmonton, Alberta, Canada, T6G 2H7 and the Canadian Bacterial Diseases Network. All other authors are with Astra Research Center Boston, 128 Sidney Street, Cambridge, MA, 02139. Putative identifications, sequence alignments, and name and sequence search capabilities are available at ARCB's World Wide Web site. (URL: <http://www.astro-boston.com/hpylori/>.)
 Location/Qualifiers
 1. 1643831
 /organism="Helicobacter pylori J99"
 /mol_type="genomic DNA"
 /strain="J99"
 /db_xref="taxon:85963"
 complement(51..467)
 /gene="nusB"
 /locus_tag="jhp_0001"
 complement(51..467)
 /gene="nusB"
 /locus_tag="jhp_0001"
 /note="similar to H. pylori 26695 gene HP0001"
 /codon_start=1
 /transl_table=11
 /product="TRANSCRIPTION TERMINATION"
 /protein_id="AAD05585.1"
 /db_xref="GI:4154504"
 /translation="MATRTQARGAVVELLYAFESGNEEIKKIASMLEEKKIKNNQLA
 FALSFLNGVLEKINEIDALIEPHLKQWDFKRLGSMKAILRLGAYEIGFTPTQNP111
 NECIELGKLYAEPNTPKFLNALDLSLKKLAQKPLI"
 complement(469..939)
 /gene="ribE"
 /locus_tag="jhp_0002"
 complement(469..939)
 /gene="ribE"
 /locus_tag="jhp_0002"
 /note="similar to H. pylori 26695 gene HP0002"
 /codon_start=1
 /transl_table=11
 /product="RIBOFLAVIN SYNTHASE BETA CHAIN"
 /protein_id="AAD05586.1"
 /db_xref="GI:4154505"
 /translation="MQIEGKQLQNEKIAITLRSNHIITRLKEGANDCFKRGHG
 DEELLDLVLPVGAELPLIDKLSGKYGCVLGLAIRGGTFPDYVSAEATKGIA
 SAMLKYSMPVSGVLTNDIEQAIERAGSKAGNKGFEAMSTLIELSLCQTLDK"
 complement(949..1779)
 /gene="kdsA"
 /locus_tag="jhp_0003"
 complement(949..1779)
 /gene="kdsA"
 /locus_tag="jhp_0003"

/note="similar to H. pylori 26695 gene HP0003"
 /codon_start=1
 /transl_table=11
 /product="3-deoxy-D-manno-octulosonic acid 8-phosphate
 synthase"
 /protein_id="AAD05587.1"
 /db_xref="GI:4154506"
 /translation="MKTENTKTPKPVLIAGPCVIESLENLRSIAIKLOPLANNERLDF
 YFKASFDKANRTSLSEYRGPGLKGLMEQTIKDERGYKILTDVHESYQASVAAKVAD
 ILQIPALPCRTQDLIVSYQTNIAIYNIKKGQPMNPKMQYSLVALKLTRDSSIQSPY
 TALKNQWYLCERGSSFGYGNLVDMRSLSKIMREFAPVIFDATHSYQMPGANGKSSG
 DSSPPILPRAAAGVIGIDGLPAETHIDPKNALSDGANMLKPDLEHLVTDMLKIQNLF
 "
 complement(1766..2431)
 /gene="icfA"
 /locus_tag="jhp_0004"
 complement(1766..2431)
 /gene="icfA"
 /locus_tag="jhp_0004"
 /note="similar to H. pylori 26695 gene HP0004"
 /codon_start=1
 /transl_table=11
 /product="Carbonic anhydrase"
 /protein_id="AAD05588.1"
 /db_xref="GI:4154507"
 /translation="MEAFIAGLEPOENEYEEPEKLYESLTKQKPHTLFISCVDSRVV
 PNLTGTPQGLYVIRNNGVIPPKTSYKSLSTIASVEYIAHGVQNLIIICGSDC
 GACSIHLIDETTKAKTPYANMTQFLEPIKEKLNHPQFNHFAKRSWLTRELNAR
 LQANLLSYDFIQERVINNELKIFGWHYIETRIYNNFESHFPEPIETIKQRIH
 ENF"
 2552..3235
 /gene="pyrF"
 /locus_tag="jhp_0005"
 2552..3235
 /gene="pyrF"
 /locus_tag="jhp_0005"
 /note="similar to H. pylori 26695 gene HP0005"
 /codon_start=1
 /transl_table=11
 /product="OROTIDINE 5'-PHOSPHATE DECARBOXYLASE"
 /protein_id="AAD05589.1"
 /db_xref="GI:4154508"
 /translation="MQLCVALDLEKEDNLSLLQELKGLDLWAKVGLRSFIRDGATFL
 DEIRKIDGNPKIFLDKLYDIPYMANAALECAKLDIMLTVHLSSAKALTILMORL
 NALKRPLIMGVLSALTSFSEBEFLMVYNAPLQAITLSAIGKESIDGVGVCSFSL
 AIKEALGKGLTLTPGIRLDKNKEDQSRVANAKAQNLSDFIVVGRPIYQAKEPRE
 VLELLKDC"
 3235..4065
 /gene="panC"
 /locus_tag="jhp_0006"
 3235..4065
 /gene="panC"
 /locus_tag="jhp_0006"
 /note="similar to H. pylori 26695 gene HP0006"
 /codon_start=1
 /transl_table=11
 /product="PANTOATE--BETA-ALANINE LIGASE"
 /protein_id="AAD05590.1"
 /db_xref="GI:4154509"
 /translation="MRALETIATLRYRKSLSKESVGFPTMGALHRGHOSLIERSLKE
 NSHTIVSVFNPTQGANEDFSAYRPLEKDILALCEKGVNAVFAPKIGEMYPYEAQ
 RLKLYAPAFSLGLEGVRKGVGVVQLRFLHNTPTTRAYFGKQDAQQLIIQHLL
 VQDLILDIEIAPCEIVRSDNLAISRNVCLNATERKQALAIKPALEIKQQAIDREK
 ACEKLLGLLEILKNLEVDYLECCNHNKLEPLKTIEPANTLVLAARVGVKTRLLDLWV
 "
 complement(5436..7443)
 /locus_tag="jhp_0007"
 complement(join(5436..7413,7415..7443))
 /locus_tag="jhp_0007"
 /note="similar to TIGR HP0009; expression of this protein
 is predicted to be regulated by slipped-strand repair at
 the (CT) dinucleotide repeat; length of the repeat has

Query Match		36.1%; Score 1040.8; DB 15; Length 110000;
Best Local Similarity		65.3%; Pred. No. 2.9e-313;
Matches 1681; Conservative		1; Mismatches 853; Indels 39; Gaps 4;
QY	67	TAATRCCTATTATTAATTTTTTAATAATTTACTT-ATTATCATATATAATAATATTATTACT 125
DB	75600	TAATACCTAATTTCAATTTAAATAATAATTTAGTTTAATAACCGTCTCTGTTAACTTTAGTAAA 75541
QY	126	TATATTTAAAAAGTTAATAAAAAAGTAACGAAATTTAGGACTATAATCCCAATTGCTTTTAAAA 185
DB	75540	TCAAAAAGATTGCTACAATTTACTCCAACTTATGTTGGTATGCTCTTCAAGGAATAACAC 75481
QY	186	TTTAAACACAGAGTAATAGGTGAACTCACACCCAAAGAGCAAGAAAGTTCTTGTTAT 245
DB	75480	TTTAAAGATAGGAAATGAGATGAACTCACCCCAAAAGAGTTAGATTAAGTTGATGCTCC 75421
QY	246	ATTATGGCGGAGTGGCTAGAAAGCGCAAGCAGAGCGGCTTAAAGCTCAACCAACCG 305
DB	75420	ACTAGCTGGAGAATTTAGCTAGGAACCGCAAGAAAGGCAATTAAGCTTTAATCTATGTGG 75361
QY	306	AAGCCATTGCTTACATTTAGTCCCATATTATGGACGAAGCGCGCTGGAAAAAAGAACCG 365
DB	75360	AAGCGGTAGCTTTGATTTAGTCCCATATTATGGAAAGCGAGAGCTGGTAAAGAGCTG 75301
QY	366	TTGCCAGCTTATGGAAGAGTGCAATGCACTTTTGAAGAAAAAGAGTAATGCCCGGG 425
DB	75300	CGGCTGAATTTGATGCAAGAGGGGCGCACTTTTAAAAACCGGATGATGTGATGGTGTG 75241
QY	426	TGGGTAAATATGTTCCCGATCTAGGTGTAGAGCCACCTTTCTGATGGTACGAACCTTG 485
DB	75240	TGGCAAGCATGATCCATGAAGTGGGTATTTGAAGCGATGTTTCTGTAGGGACCAAACTCG 75181
QY	486	TAACTGTGAATTTGGCCCATCGAAACAGATGAGCACTTCAAAGCGGGCGAAGTGAATTTG 545
DB	75180	TAACTGTGCAATCCCTATTGAGGCTAATGTAATTTGGTTCCTGGTGAATG--TTCT 75124
QY	546	GTTGCGATAAGACATCGAGCTCAATGAGGCAAGAAAGTAACCGCACTTGAGGTACTA 605
DB	75123	TAAAAATGAAGACATCACTATCAAGGAAGGCAAAAAGCGGTTAGCGTGAAGTTAAAA 75064
QY	606	ATGAAGGCTTAATCTTGCATGTGGGTAGGCATTTCCACTTCTTTGAAAGTAACAAG 665
DB	75063	ACGTGGGCGACAGACCGGTTCAAAATCGGTTTCACTTCCATTTCTTTGAAAGTAATGAT 75004
QY	666	CACATAAATTCGATCGTGAAGAAAGCTATGCAAGCGCTAGATATTCCTCTGCAACA 725
DB	75003	GCCTAGACTTTGACAGAGAAAAAACTTTCGGCAAGCGCTTAGACATTCGAGCGGACAG 74944
QY	726	CGCTACGCAATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTA 785
DB	74943	CGGTAAAGTTTGAAGCTTGGCGAGAAAGAAATCGTAGAATTTGATGACATTTGTTGGCAACA 74884
QY	786	AAAAAGTGAATGGCAATGACGGCTTTGTGAATAACATCGGGATGAACGCCATAACATA 845
DB	74883	GAAGAACTTTTGGATTTAACGCAATTTGGTTGATAGGCAAGCAACAACGAAGCAAAAAA 74824
QY	846	AAGCGCTTGAACAGCGAAATCTACGGAATTT----- 877
DB	74823	TTGCTTTTACACAGACTAAAGAGCGTGGTTTTCATGGTGTAAAAAGCGATGACAACTATG 74764
QY	878	----ATCAAGTAAGGAGACTCCATGAAA--ATCAAAAAACAAGATATGTAAATACCT 930
DB	74763	TAAAAACAAATTAAGAGGTAAAGAAATGAAGAAATTAAGCAAGAAAGATATGTTCTATGT 74704
QY	931	ACGGAACCAACCAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAG 990
DB	74703	ATGGCCCTACACAGCGGATAAAGTGAATTTGGCGGATACAGATTTGATCGTGAAGTAG 74644
QY	991	AACATGACTATACCACTATGCGGAGAACTTAAATTTGGCGGGTAAACTATCCGTG 1050
DB	74643	AACATGACTATACCACTTTATGCGGAGAGCTTTAAATTCGGTGGCGGTAAAAACCTTAAGAG 74584
QY	1051	AGGATATGGGTGAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTTAGTCACTAACG 1110
DB	74583	AAGCATGAGCAATCCAAACACCCCTAGCAAAAGAAACTGGATTTGATCATCACTAACG 74524
QY	1111	CGATGATTATCGACTACACCGGATTTTACAAAGCCGACATTTGGGATTTAAAAACGGCAAAA 1170
DB	74523	CTTTAATCGTGGATTTACACCGGTATTTTAAAGCGGATTTGGTATTTAAAGACGGCAAAA 74464
QY	1171	TCCATGGCAATTTGGCAAGGACGAGAAAACAAGGACATGCAAGATGGCGTAAGCCCTCATATGG 1230
DB	74463	TCGCTGGCAATTTGGTAAAGCGGTAAACAAGACATGCAAGATGGCGTTAAAAACAATCTTA 74404
QY	1231	TCGTGGGTGGGACACAGAGCACTAGCAGGGGAAGGTATGATTATTACCGCTGGGGGAA 1290
DB	74403	CGGTGGGTCTGCTACTAGCGCTAGCTGGTGAAGGCTTAATCGTAAACGGCTGGTGGTA 74344
QY	1291	TCGATTTACACACACCCACTTCTTCTTCCACAACAATTCCTTACCCTCTAGCCCAATGGCG 1350
DB	74343	TTGACACACACATCCACTTCAATTTACCCCAACAATCCCTACAGCTTTTGCAGCGGTG 74284
QY	1351	TTACAACCATGTTTGGAGCGGCAAGGTCTGTAGATGGCAAGATGGGACTTACTATCA 1410
DB	74283	TAAACAACCATGATTGGTGGCGAAACGGGTCTGCTGTAGTGGCACTAATGCACTACTATCA 74224
QY	1411	CTCCGGGCAATTTGGAACCTTTGACCGCATGTTGGGCGCAGCAGAGATTTCTATGAATG 1470
DB	74223	CTCCAGGTAGAAGAAACTTTAAATGGATGCTCAGAGCGGCTGAAGAAATATTCTTATGAAT 74164
QY	1471	TGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTTGTAACAAGTAGAAG 1530
DB	74163	TAGGTTTCTTAGCTAAGGTAACGTTCTTAACGACGAGCTTAGCGGATCAAAATTTGAAG 74104
QY	1531	CGGCGCGAATTTGGTTTAAATTTGCAATGAGACTTGGGGCAACAACCAAGTGGGATCGATC 1590
DB	74103	CGGTGGCATTTGGCTTTAAATCCACGAGACTTGGGGCAACCACTCTCTTCTGCAATCAATC 74044
QY	1591	ACTGCTGAGCGTGGCAGATGAATGAGTGTGCAAGTTTGTATTCACACCGATACAGTCA 1650
DB	74043	ATGCGTTAGACGTTGCGGCAAAATACGATGTCAAGTTCGCTATCCACACAGACACTTTGA 73984
QY	1651	ATGAGCGAGTTTATGTAGATGACCCCTAAATGCAATGAACGGCGCGCATCCATGCGCT 1710
DB	73983	ATGAAGCGGTTTGGTGGAAAGACATATGGCAGCTATTGCCGACGCACTATGCACTT 73924
QY	1711	ACCAATTTGAGGAGCGGTGGAGGACACTCACTGATTTATCACCATGGCAGGCGGAGC 1770
DB	73923	TCCACACTGAAGCGCTGGTGGCGACACGCTCTGACATTTAAAGTGGCGCGTGAAC 73864
QY	1771	TCAATTTCTACCTCTCCACCAACCCCACTATTTCCTTATACCATTAATACGGTTGCGAG 1830
DB	73863	ACAACATCTACCCGCTTCCACTAACCCCACTATCCCTTTCACTGTGAATACAGAAAGCCG 73804
QY	1831	HACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGATTAC 1890
DB	73803	AACACATGACATGCTCATGTTGGTGTGCCCACTTTGGATAAAAAGCAFTTAAAGAGATGTCC 73744
QY	1891	AAATTTCTCAAGCGCTATCCCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCATGATA 1950
DB	73743	AGTTGCTGATTTCAAGGATCCGCCCTCAAAACCAATTGGCGCTGAAGACACTTTGCAATGACA 73684
QY	1951	TGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGGCGAAGTGA 2010
DB	73683	TGGGGATTTTCTCAATCACAGTCTTCTGACTCTCAAGCGATGGGTGCTGTGGTGAAGTTA 73624
QY	2011	TTCTCGAACTTTGGCAGACTCGGATTAAGAAATAAAAAAGAAATTTGGTAAAGTCTTCTGAAG 2070
DB	73623	TCACCTAGGACTTTGGCAACAGCTGCAAAAAAATAAAAAAGAAATTTGGCCGCTTGAAGAAG 73564
QY	2071	ATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAATACACTATCAACC 2130
DB	73563	AAAAAGCGATTAACGACAACTTCAGGATCAACGCTACTTGTCTAANTACACCACTTAACC 73504
QY	2131	CGGCTTTGACCCACCGCGTGAAGAGTATATCGGCTCTGTGGAAAGGGCAAGATCGCG 2190

Db	73503	CAGCGATCGCTCATGGGATTAGCGAGTATGTAGGTTCTGTAGAAAGTGGCAAAGTGGCTG	73444
Qy	2191	ACTTGGTGGTGTGAATCTCTGCTTTTGGGTAAACCCAAATCGTATCAAGCGC	2250
Db	73443	ACTTGGTATTGTGGAGTCCAGCATCTTTGGCGTGAACCCAAACATGATCAAGCGC	73384
Qy	2251	GTATGGTGGTCTTCTCGAAATGGCGATTCTAAGCGCTGTGCGCACTCCCAACCGG	2310
Db	73383	GATTCATGGCTTAAGCAAAATGGCGATGGAAGCGTCTTATCGCTACCCACACACCGG	73324
Qy	2311	TTTATTACCGGAAATGTTTGGGCATCAGCGCAAGCGGAATTTGACACGACATCACTT	2370
Db	73323	TTTATTACAGAGAATGTTGCTCTCACCATGTTAAGCTTAATACATGATCAAAACATCACTT	73264
Qy	2371	TGTTTCAAAAGTCGCTATGAATGCGTGAAGAAAGCTGGCTTAGAGCGCCAG	2430
Db	73263	TTGTGTCGCAAGCGCTTATGACAAAGCATTAAGAAGAAATTAGGGCTTGAAGAACAAG	73204
Qy	2431	TTCTACCGGTCAAAACCTGCGTACATCAACCAAGAAGACTTCAAGTTCAACGACAAA	2490
Db	73203	TGTTGCGCGTTAAATAATTGCAGAACATCACTAAAAAGACATGCAATTCATGACACTA	73144
Qy	2491	CGGCAAAATCACCGTCGATCCGAAACCTTCGAGGCTTTGTAGATGGCAAACTCTGCA	2550
Db	73143	CGCTCATATGAAGTCAATCTGAACTTACATGTGTTCTGTGATGGCAAGAGTCA	73084
Qy	2551	CCTCTAAACCCACTCGCAAGTGCCTCTTAGCCAGCGCTACACTTCTCTTAGG	2604
Db	73083	CTTCTAAACCAAGCTAATAAAGTGAGCTGGCTCAACTCTTTAGCATTTTCTTAGG	73030
RESULT 14			
WPCOMMENT			
Sequence split into 17 fragments LOCUS AE000511 Accession AE000511			
Fragment Name Begin End			
AE000511_00 1 110000			
AE000511_01 100001 210000			
AE000511_02 200001 310000			
AE000511_03 300001 410000			
AE000511_04 400001 510000			
AE000511_05 500001 610000			
AE000511_06 600001 710000			
AE000511_07 700001 810000			
AE000511_08 800001 910000			
AE000511_09 900001 1010000			
AE000511_10 1000001 1110000			
AE000511_11 1100001 1210000			
AE000511_12 1200001 1310000			
AE000511_13 1300001 1410000			
AE000511_14 1400001 1510000			
AE000511_15 1500001 1610000			
AE000511_16 1600001 1667867			
LOCUS	AE000511	1667867 bp	DNA circular BCT 27-DEC-2005
DEFINITION	Helicobacter pylori 26695, complete genome.		
ACCESSION	AE000511	AE000523-AE000656	
VERSION	AE000511.1	GI:6626253	
KEYWORDS			
SOURCE	Helicobacter pylori 26695		
ORGANISM	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.		
REFERENCE	1 (bases 1 to 1667867)		
AUTHORS	Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.		
TITLE	The complete genome sequence of the gastric pathogen Helicobacter pylori		

JOURNAL	Nature 388 (6642), 539-547 (1997)
PUBMED	9252185
REFERENCE	2 (bases 1 to 1667867)
AUTHORS	Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Smith, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.
TITLE	Direct Submission
JOURNAL	Submitted (06-AUG-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
REFERENCE	3 (bases 1 to 1667867)
AUTHORS	White, O.
TITLE	Direct Submission
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
COMMENT	On or before Dec 27, 2005 this sequence version replaced gi:2313077, gi:2313090, gi:2313102, gi:2313116, gi:2313131, gi:2313137, gi:2313152, gi:2313162, gi:2313173, gi:2313184, gi:2313196, gi:2313205, gi:2313217, gi:2313230, gi:2313247, gi:2313263, gi:2313275, gi:2313287, gi:2313299, gi:2313310, gi:2313321, gi:2313337, gi:2313349, gi:2313363, gi:2313377, gi:2313391, gi:2313403, gi:2313417, gi:2313430, gi:2313451, gi:2313461, gi:2313475, gi:2313485, gi:2313505, gi:2313581, gi:2313593, gi:2313602, gi:2313616, gi:2313628, gi:2313641, gi:2313646, gi:2313663, gi:2313672, gi:2313686, gi:2313703, gi:2313712, gi:2313721, gi:2313730, gi:2313736, gi:2313747, gi:2313759, gi:2313778, gi:2313792, gi:2313802, gi:2313812, gi:2313824, gi:2313834, gi:2313845, gi:2313854, gi:2313869, gi:2313880, gi:2313895, gi:2313907, gi:2313918, gi:2313929, gi:2313944, gi:2313957, gi:2313969, gi:2313982, gi:2314007, gi:2314019, gi:2314028, gi:2314042, gi:2314051, gi:2314060, gi:2314075, gi:2314086, gi:2314103, gi:2314115, gi:2314125, gi:2314139, gi:2314150, gi:2314160, gi:2314173, gi:2314188, gi:2314200, gi:2314216, gi:2314230, gi:2314242, gi:2314256, gi:2314268, gi:2314276, gi:2314293, gi:2314301, gi:2314317, gi:2314327, gi:2314340, gi:2314349, gi:2314360, gi:2314373, gi:2314386, gi:2314400, gi:2314409, gi:2314421, gi:2314436, gi:2314452, gi:2314489, gi:2314502, gi:2314517, gi:2314536, gi:2314547, gi:2314560, gi:2314572, gi:2314581, gi:2314587, gi:2314598, gi:2314609, gi:2314623, gi:2314635, gi:2314645, gi:2314670, gi:2314687, gi:2314700, gi:2314708, gi:2314720, gi:2314733, gi:2314743, gi:2314757, gi:2314771.
FEATURES	Location/Qualifiers
source	1. 1667867
	/organism="Helicobacter pylori 26695"
	/mol_type="genomic DNA"
	/strain="26695"
gene	/db_xref="taxon:85962"
	complement (217..633)
CDS	/locus_tag="HP_0001"
	complement (217..633)
	/locus_tag="HP_0001"
	/note="hypothetical protein; identified by Genemark; putative"
	/codon_start=1
	/transl_table=11
	/product="predicted coding region HP0001"
	/protein_id="AA07074.1"
	/db_xref="GI:2313078"
	/translation="MATRTQARGAVVELLYAFESNGNEIKTKASSMLEKKIKNNOLA FALSFLNGLEKINEIDALIEPHLKDWDFKRLGSMEKAILRGAYEIGTPTQNPILII NECTELGKLYAEPTNPFLNALIDSLSKLTQKPLN"
gene	complement (635..1105)
	/locus_tag="HP_0002"
CDS	complement (635..1105)
	/locus_tag="HP_0002"
	/note="similar to GB:L42023 SP:P45149 PID:1007275"

Db 77319 TTGCTTTACACAGAGCTAAAGAGCGTGGTTCATGGCGCTAAAAAGCGATGACAACTATG 77260
Qy 877 T---ATCAAGTAAGAGAGCTCCCATGAA---ATCAAAAAACAGAAATATGTAATACCT 930
Db 77259 TAAAAACAATTAAGAGATGAAGAAATGAAGAAATAGCAGAAAAAGAAATATGTTCTTATGT 77200
Qy 931 ACGGACCCACCAAAAGGCGATAAAGTGGCGTTAGGAGATACCGATCTTTGGGCAGAGTAG 990
Db 77199 ATGGCCCTACTACAGCGATTAAGTGAATTTGGCGGATACAGACTTGATCGCTGAAGTAG 77140
Qy 991 AACATGACTATACCACTATCGGCAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTG 1050
Db 77139 AACATGACTACCACTTATGCGCAAGAGCTTAAATTCGGTGGCGGTAAAAACCCCTGAGAG 77080
Qy 1051 AGGTATGGTTCAGAGCAATAGCCCTGATGAAGAACACCCCTAGATTTAGTCTATCACTAACG 1110
Db 77079 AAGGCATGAGCAATCCAAACAACTTACCAAGAGAAATTTGGATTAATCACTACCTAACG 77020
Qy 1111 CGATGATTATCGACTACACCGGGATTTACAAGCGGACATTTGGGATTAAGAAACGCAAAA 1170
Db 77019 CTTTAAATCGTGATTAACCGGATTTTATAAGCGGATTTGGTATTAAAGATGGCAAAA 76960
Qy 1171 TCCATGGCATTTGGCAAGCGAGAAACAGGACATGCAAGATGGCGTAAAGCCCTCATATGG 1230
Db 76959 TCGCTGGCATTTGTAAGCGGTAAACAAGACATGCAAGATGGCGTTAAACAACATCTTA 76900
Qy 1231 TCGTGGGTGGGCAAGAAGCACTAGCAGGGGAAGTATGATTTATACCGCTGGGGAA 1290
Db 76899 GCGTAGGTCCTGCTACTGAAAGCCTTAGCCGGTGAAGGTTTGATCGTAACTGCTGGTGA 76840
Qy 1291 TCGATTCACACACCACTTCTCTTCTCCAGAAATTTCCCTACCGCTTAGCCATGGCG 1350
Db 76839 TTGACACACATCCACTTCATTTTACCCCAACAAATCCCTACAGCTTTTTCAGAGCGGTG 76780
Qy 1351 TTACAACATGTTTGGAGGGGACAGCTCTGTAGATGGCAACAAATGCGACTACTATCA 1410
Db 76779 TAACAACATGATTTGGTGGGAACTGCTGCTGTGATGGCACTTAATGCGACTACTATCA 76720
Qy 1411 CTCGGGCAATGAACTTGCACGCGATTTTGGCGCAGCAGAGAGTATTTCTATGAATG 1470
Db 76719 CTCAGGCAAGAATAATTAATGGATGCTCAGAGCGGCTGAAGATATTTCTATGAAT 76660
Qy 1471 TGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAAACAACTTGTAGAACAACTAGAG 1530
Db 76659 TAGGTTTCTGGCTAAAGGTAAACGCTTTTACGACGCGGAGCTTAGCCGATCAAAATTTGAAG 76600
Qy 1531 CGGCGCGATTTGTTTAAATTTGATGATGAACTGGGCAACACCAAGTCGATCGATC 1590
Db 76599 CTGGTGGATTTGGCTTTTAAATTCACGAAGACTGGGGCAACCACTCTTCTGCAATCAATC 76540
Qy 1591 ACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCA 1650
Db 76539 ATCGTTAGATTTGACAGACAAATACGATGTGCAAGTCGCTATCCACACAGACTTTGA 76480
Qy 1651 ATGAGCGGTTGCGTGGAAAGACACTATGGCAGCTATTTGGCGGACGCACTATGCACACT 76420
Db 76479 ATGAGCGGTTGCGTGGAAAGACACTATGGCAGCTATTTGGCGGACGCACTATGCACACT 76420
Qy 1711 ACCACATTTAGGAGCGGTTGGAGGACACTCACCCTGATTTATCACCATGGCAGCGAGC 1770
Db 76419 TCCACACTGAAGTGTGGCGGGGACACGCTCTGATATTTAATTAAGATGAGTGTGTAAC 76360
Qy 1771 TCAATATTTACCTCTCTCCACACCCCACTATTCCTATATCACTTAATTAAGCTTGCAG 1830
Db 76359 ACAACATTTTCCCGCTTCCACTTAACCCCACTATTCCTTTCATGTAATACAGAGCAG 76300
Qy 1831 AACACTTAGACATGCTCATGACATGCCACCACTTAGCAACAACGATCCGGGAGATTAC 1890
Db 76299 AACACATGACATGCTTATGTTGGTGGCCACCACTTGGATTAAGACATTAAGAGATGTT 76240
Qy 1891 AATTTTCTMAAGCGGTATCCGGCCCGGCTCTATCGCGGCTGAAGATGTGCTCATGATA 1950
Db 76239 AGTTCGTGATTTCAAGGATCCGGCCCTCAAAACCAATTTGGGCTGAAGACACTTTTGCATGACA 76180

Qy 1951 TGGGTGTGATCGCGATGACAACTCGGATTCGCAAGCAATGGGCGGTGCGAGGCAAGTGA 2010
Db 76179 TGGGGATTTTCTCAATCACAGCTCTGACTCTCAAGCTATGGGTGCTGGTGAAGTTA 76120
Qy 2011 TTCTCTCGAACTTGGGAGAGTGGGATGAAGATAAATAAAGAAATTTGGTAAAGCTTCTCTGAAG 2070
Db 76119 TCACTAGAACTTGGCAACAGCTGACAAAAAATAAAGAAATTTGGCCGCTTGAAGAAG 76060
Qy 2071 ATGGCAAGATAACGATTAATTTCCGCAATTAAGCGCTACATCTCAATAACATATCAACC 2130
Db 76059 AAAAAAGCGCATACGCAACTTTCAGGATCAACAGCTACTTGTCTAAATACACCATTAACC 76000
Qy 2131 CGGCTTTGACCCACCGCTGAGGAGTATATCGGCTCTGTGGAGAGGGCAAGATCGCG 2190
Db 75999 CAGCGATCGCTCATGGGATTAGCGAGTATGTAGGTTCTGTAGAGTGGGCAAGTGGCTG 75940
Qy 2191 ACTTGGTGTGGTGAATCTGCTCTTTTGGGTAAAAACCCAAAAATCGTATCAAGGGCG 2250
Db 75939 ACTTGGTATTTGGAGTCCGCAATTTTGGGCTAAAAACCCAAATGATCATCAAGGGCG 75880
Qy 2251 GTATGGTGTCTTCTCTGAAATGGGCGATTTCTAACCGCTCTGTGCCACCTCCCAACCGG 2310
Db 75879 GGTTCATTTGGTTGAGTCAATGGGTGACGCGAACGCTTCTATCCCTACCCCAACCGG 75820
Qy 2311 TTTATTAACGGGAAATGTTTGGGCTATCAGCGAAGCGGAAATTTTGACACAGCATCACTT 2370
Db 75819 TTTATTAACAGAGAAATGTTTGGCTCATCATGTTAAAGCCAAATACGATGCAAAACATCACTT 75760
Qy 2371 TTGTTTCCAAAGTCCGCTATGAAATGGGTGCAAGAAAGCTGGGCTTAGAGCGGCAAG 2430
Db 75759 TTGTTCTCAAGCGGCTTATGACAAAGGCAATTAAGAAAGATTTAGGGCTTTGAAGACAAG 75700
Qy 2431 TTCTACCGCTCAAAAACCTGCCGTAAACATCAACCAAGAAAGACTTCAAGTTTCAACGCAAAA 2490
Db 75699 TGTTCGCGGTAAATAATTTGCAAAACATCACTTAAAGAAACATGCAATTTCAACGACTA 75640
Qy 2491 CGGCAAAATCACCGTCAATCGATCGAAGAACTTCAGAGTCTTTGTAGATGGCAAACTCTGCA 2550
Db 75639 CGGCTCAATTTGAAGTCAATCTGAAACTTACCATGTGTTGGTGGATGGCAAGAAAGTAA 75580
Qy 2551 CCTTAACCCACCTCGCAAGTGCCTTAGCCAGCGCTACACTTTCTTCTAGG 2604
Db 75579 CTTCTAAACCAAGCAATAAAGTAGAGCTTGGCGCAACTCTTTAGCATTTTCTAGG 75526

RESULT 15

HECUREABCD
LOCUS HECUREABCD 5100 bp DNA linear BCT 26-APR-1993
DEFINITION H.pylori urease (ureA, ureB, ureC, ureD) genes, complete cds.
ACCESSION M60398 X57132
VERSION M60398.1 GI:149007
KEYWORDS urease.
SOURCE Helicobacter pylori
ORGANISM Helicobacter pylori
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE
1 (bases 1 to 5100)
Labigne, A., Cussac, V. and Courcoux, P.
Shuttle cloning and nucleotide sequences of Helicobacter pylori
genes responsible for urease activity
J. Bacteriol. 173 (6), 1920-1931 (1991)
2001995
COMMENT Original source text: H. pylori DNA.
FEATURES
Location/Qualifiers
1..5100
/organism="Helicobacter pylori"
/mol_type="genomic DNA"
/db_xref="taxon:210"
510..1847
/gene="ureC"
510..1847
/gene="ureC"

[illegible]

ORIGIN

Query Match 35.6%; Score 1024.4; DB 15; Length 5100;
Best Local Similarity 65.3%; Pred. No. 3e-308;
Matches 1682; Conservative 1; Mismatches 852; Indels 39;

67	QY	TAATRC	TATATTAA	TTTTTAA	TAAATTA	CTT-ATT	ATCATAT	TATAT	TAATAT	TATTTACT	125
		:									
2519	Db	TAACAC	TAAATTC	ATTTTAA	AATAATA	TTAGTT	AATGA	CGCTT	CTGTT	TAATCTTAGT	2578
126	QY	TATAT	TAAAAAG	TTATAAA	AGTAAG	AAATTA	GAGCTA	TATAT	CCCAT	TGCGCTTTAAAA	185
2579	Db	TCAAAA	CATTTG	CTACAT	CACTCC	CAACCT	TGAT	TGCTT	GCCTT	CAAGGAAAAACAC	2638
186	QY	TTTAA	CACAGG	AGTTAT	TAGGT	GAAT	CTCAC	CCCAAG	AGCAAG	AAAAAGTTCTTGT	245
2639	Db	TTTAA	GAATAGG	AGAAAT	GACAT	CGAACT	CACCC	CAAGAG	TGTAG	TAAGTGTGATGCTC	2698

Qy	246	ATTATCGGGCGAGTGGCTAGAAAGCGCAACGACGAGGCGTTAAAGCTCAACCAACCCG	305
Db	2699	ACTACGCTCGGAGAAATGGCTTAAAAAACGCAAGAAAAAGSCATTAAGCTTTAACTATGTAG	2758
Qy	306	AAGCCATTGCTTACATTTAGTGCCCATATTATGACGAAGCGCGCTGGAAAAAACAACCG	365
Db	2759	AAGCAGTAGCTTTGATTTAGTGCCCATATTATGAGAAGCGAGCTGGTTAAAAAGACTG	2818
Qy	366	TTGCCCAGCTTATGGAAGAGTGCATGCACCTTTTGTGAAAAAAGATGAAGTAATGCCCCGGG	425
Db	2819	CGGCTGAATTTGATGCAAGAAGGCGCACTCTTTTAAAAACAGATGATGTGATGATGCG	2878
Qy	426	TGGGTAAATATGTTTCCCGATCTAGGTGTAGAGCCACTTTTCCCTGATGGTACGAACCTTG	485
Db	2879	TGGCAAGCATGATCCTCATGAAGTGGGTATTGAAGCGATGTTTCCCTGATGGGACTAAACCTG	2938
Qy	486	TAACTGTGAATTTGGCCCATCTGAACACAGATGAGCACATTCAAAGCGGCGAAGTGAATTTTG	545
Db	2939	TAAACGTGATACCCCTATTGAGCCAAATGGTAAAAATTAGTTCTGTGTAGTTG---TTCT	2995
Qy	546	GTTTGCATATAAGACATCGAGCTCAATCTCAGGCAAGAAAGTAAACCGCACTTTGAGGTTACTA	605
Db	2996	TAAAAAATGAGACATCACTATCAAAGAGGCAAAAAAGCGGTAGCGTGAAGTTAAAA	3055
Qy	606	ATGAAGGGCCTAAATCCTTTGCATGTGGGTAGCCATTTTCCACTTTCTTTGAAGCTTAACAAG	665
Db	3056	ATGTTGGCAGACAGACGGCTTCAATCGGCTCACACTTCCATTTCTTTGAAGTGAATAGAT	3115
Qy	666	CACATAAAATTCGATCGTGAAGAACCTNTGGCAAAACGCTTAGATATTCCTCTGGCAACA	725
Db	3116	GCCTAGACTTTGACAGAGAAAAAACTTTCGGTAAACGCTTTAGACAATTGGCAGCGGACAG	3175
Qy	726	CGCTACGCATTGGGGCAGGACAAACCGCAAGTGCAGTTGATTCTCTTTGGTGGCAGTA	785
Db	3176	CGGTAGATTGGAGCTTGGCGAAGAAAAATCCGTAGAAATTTGATTGTGACATTTGGCGGTAA	3235
Qy	786	AAAAAGTGAATGGCATGAACGGGCTTTGTGAATAACATTCGCG---ATGAACGCCA----	837
Db	3236	GAAGAACTTTGGATTTAACGCATTGGTTGATAGACAAGCAGACAACGAAGCAAAAAA	3295
Qy	838	-----TAAACA-----TGAACAGCGC-----AA-----ATCTACGGATT	876
Db	3296	TTGCTTTTACACAGACTTAAGAGCGTGGTTTTCATGGCGCTTAAAGCGCATGACAACTATG	3355
Qy	877	T---ATCAAGTAGAGGAGACTCCCATGAA--ATGAAAAACAGATAATGTAAATACCT	930
Db	3356	TAAAAACAATTAGGAGTAGAAATGAAAAAGATTAGCAGAAAAGAAATATGTTTCTATGT	3415
Qy	931	ACGGACCCCAAGGCGATAAAGTGCGCTTAGGAGATACCGATTCTTTGGGCAGAACTAG	990
Db	3416	ATGGTCTACTACAGGCGATAAGATGAGATTGGCGATACAGACTTGATCGCTGAAGTAG	3475
Qy	991	AACATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGGTAAAACTATCCGTG	1050
Db	3476	AACATGACTACACCAATTATTTGGCGAAGAGCTTAAATTCGGTGGCGGTAAAAACCCCTAAG	3535
Qy	1051	AGGGTATGGTCHAGACCAATAGCCCTGATGAAAACACCCTAGATTTAGTTAGTCACTAACG	1110
Db	3536	AAGGCATGAGCCCAATCTTAAACAACCCCTAGCAAGAAAGAGTTTGGATTTTAAATTTATCACTAACG	3595
Qy	1111	CGATGATTATCGACTACACCGGGATTTFACAAAGCCGACATTTGGGATTTAAAAAACGGCAAAA	1170
Db	3596	CTTTAATCTGTGGATTACCCGGTATTATTAAAGCGGATATTGGTATTAAAGATGGCAAAA	3655
Qy	1171	TCCATGGCATTTGGCAAGGCAAGAAACAGGACATGCAAGATGGCGTAAAGCCCTCATATGG	1230
Db	3656	TCGCTGSCATTGTTAAGGCGGTAAACAAAGACATGCAAGATGGCGTTAAAAAACAATCTTA	3715
Qy	1231	TCGTGGGTGTGGGCACAGAAGCACTAGCAGGGGAAGGTATGATTATTACCGCTGGGGGAA	1290
Db	3716	GCGTAGCTCTGTACTTGAAGCCTTAGCCGGTGAAGCTTTGATCGTAACGGCTGGTGTA	3775

Qy	TCGATTCACACACCCACTTCCTCTTCTTCCACAAACAATTTCCCTACCGCTCTAGCCAAATGGCG	1355
Db	TTGACACACACATCCACTTCATTTACCCCAACAAATCCCTACAGCTTTTTCGAAGCGGTG	3835
Qy	TTACAACCATGTTTGGAGGGGACAGGTCCTGTAGATGCGACGAAATGCGACTACTATCA	1410
Db	TACAACCATGATTTGGTGGTGGAAACCGGTCCTGTGATGGCACTAATGCGACTACTATCA	3895
Qy	CTCGGCGCAATGGAACTTGCCACCGCATGTTGGCGCGCAGCAGAGAGATATTCCTATGAATG	1470
Db	CTCAGCGCAGAGAAATTTAAATATGGATGCTCAGAGCGGCTGAAGATATTCCTATGAATTT	3955
Qy	TGGGCTTTTGGGCAAGGCAATAGCTCTTAGCAAAAAACAACCTGTAGAACAAAGTAGAAG	1530
Db	TAGGTTTCTTGGCTTAAAGGTAAACGCTTCTTAACGATGCGAGCTTAGCCGATCAAAATTTGAAG	4015
Qy	CGGGCGGATTTGGTTTTAAATTCGATGAAGCACTGGGGGCAAAACCAACGATGGCGATCGATC	1590
Db	CCGGTCGATTTGGCTTTTAAATTTACGAAGACTTGGGGCACCACTCTCTTCTGCAATCAATC	4075
Qy	ACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATTCACACCGATACAGTCA	1650
Db	ATGGTTAGATGTTTGGGCAAAATACGATGTGCAAGTCGCTATTCACACAGACACTTTTGA	4135
Qy	ATGAGCAGGTTATGTAGTAGACACCCATAAATCGAATGAACGGGCGGCGCATCCATGCGCT	1710
Db	ATGAAGCGGTTGTGTAGAAGACACTATGCTGCTATTGTGTGACGCACTATGCGACACTT	4195
Qy	ACCACATTTAGGAGCGGGTGGAGGACACTCACTGATGTTATCACCATGGCAGGCGAGC	1770
Db	TCACACACTGAAGCGCTTGGCGGGGACACGCTCCTGATATTTATTAAGTAGCGGCTGAAAC	4255
Qy	TCAATATTTACCTCTCCACCAACCCCACTATTCCTATACCATTAATACGGTTGCGAG	1830
Db	ACAACATTTCTCGGCTTCCACTAAACCCACCACTCCCTTTCACCGTGAATACAGAAGCAG	4315
Qy	AACACTTTAGACATGCTCATGACATGCCACCACTAGACAAACGCATCCGCGAGGATTTAC	1890
Db	AGCACATGGACATGCTTATGTTGTGCGCACCACTTGGATTAAGCACTTAAAGAAAGATGTTT	4375
Qy	AATTTTCTCAAAGCGTATCCGGCCCGGCTCTATCCGGCTGAAGATGTGCTCCATGATA	1950
Db	AGTTCGCTGATTCAAGGATCCGGCTCTCAAAACCAATTCGGCTGAAGACACTTTTGCAATGACA	4435
Qy	TGGGTGTATCGCGATGACAAAGCTTCGATTCGCAAGCAATGGGCGCTGCAGGCGAAGTCA	2010
Db	TGGGATTTTCTCAATCAACAGTTCTGACTCTCAAGCGATGGGCGGTGGGTGAAGTTTA	4495
Qy	TTCTCTCGAACTTGGCAGACTCGGATAAGAAATAAAAAGAAATTTGGTAAAGTTCCTGAAG	2070
Db	TCACTAGAACTTGGCAAAACAGCTGACAAAAACAAGAAAGAAATTTGGCCGCTTGAAGAAAG	4555
Qy	ATGGCAAGATACGATAATTTCCGCAATTAAGCGCTACATCTCCAATACACTATCAACC	2130
Db	AAAAAGCGGATAACGAACTTCAGGATCAAAACGCTACTTGTCTAAATACCACTTTAACCC	4615
Qy	CCGCTTTGACCCACGCGTCGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGG	2190
Db	CAGCGATCGCTCATGGGATTTAGCGAGTATGTAGGTTCAGTAGAAGTGGGCAAGTGGCTG	4675
Qy	ACTTGGTGTGTGGAATCTCGCTTTTTTTGGCGTAAAAACCCAAAATCGTGATCAAGGCG	2250
Db	ACTTGGTATTTGGAGTCCAGCATTTCTTTGGCGTGAAACCCCAACATGATCAAAAGCGG	4735
Qy	GTATGGTGTCTTCTCTGAAATGGGCGATTTCTAAACGCGTCTGTGCCACTCCCCAAACCGG	2310
Db	GATTCATTTGGTTTAAAGCAAAATGGGCGGATGCGAAACGCTTCTATCCCTACCCCAACCGG	4795
Qy	TTTATTTACCGCAAAATGTTTGGGCATCACGGCAAGCGGAAATTTTGACACAGCATCACTT	2370
Db	TTTATTTACAGAGAAATGTTTCGCTCATCATGTTTAAAGCTAAATACGATGCAAAACATCACTT	4855
Qy	TTGTTTTCCAAGTTCGCTATGAAATTTGGCGTGAAGAAAGAAAGCTTGGGCTTAGAGCGCCAG	2430

Accession	Sequence	Length
Db	TTTGTTCTCAAGCGGCTTATGACAAAGGCAATTAAGAAGATTAGACTTTGAAGACAAG	4915
Qy	TTCTACCGGTCAAAAACCTCCGCTTAACATCACCAGAAAGACATTCAGTTTCAACGACAAA	2490
Db	TGTTGCCGGTAAAAAATTGCAGAAATATCACTAAAAAAGACATGCAATTTCAACGACACTA	4975
Qy	CGGCAAAATATCACCCTCGATCCGAAAACCTTCGAGGTCTTTGTAGATGCGAAACTCTGCA	2550
Db	CTGCTCACATTTGAAGTCAATCCTGAAACCTTACCATGTGTTCGTGGATGCGCAAGAAGTAA	5035
Qy	CCTCTAAAACCCACCTCGCAAGTGCCTCTAGCCCGCGCTACACTTTCCTCTAGG	2604
Db	CTTCTAAACCGACCAATAAAGTGAGCTTCGGCGCACTCTTTAGCAATTTTCTAGG	5089

Search completed: August 10, 2006, 06:13:57
Job time : 16259 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2006, 01:12:00 ; Search time 1748 Seconds
(without alignments)
11499.424 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2880.6
Sequence: 1 rgragatttccarctt.....aaaaaagtagaacacagg 2883

Scoring table: IDENTITY_NUC

Gapop 3.0 , Gapext 3.0

Searched: 524920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_8.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001s.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002s.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003s.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004s.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*
- 15: Geneseqn2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2880.6	100.0	2883	7	ADJ58237
2	2228	77.3	2452	7	ADJ58249
3	2158.6	74.9	2405	7	ADJ58240
4	2158.6	74.9	2407	7	ADJ58246
5	1948	67.6	2183	7	ADJ58243
6	1134.4	39.4	8407	12	ADQ37847
7	1021.4	35.5	2735	2	AAT44351
8	1008.8	35.0	2619	2	AAQ90180
9	1008.8	35.0	2619	2	AAT45680
10	1007.2	35.0	2619	2	AAQ75319
11	1005.8	34.9	4824	5	AB551130
12	1005.8	34.9	4824	5	AB551112
13	1004.8	34.9	2767	2	AAQ12485
14	940.8	32.7	1710	12	ADQ37850
15	902.4	31.3	1719	8	ABA00816
16	885.8	30.8	1815	2	AAAI4230
17	884.8	30.7	1710	8	ACA34557
18	881.8	30.6	1707	13	ADU05358

19	881.8	30.6	1707	13	ADU05359
20	880	30.5	1710	9	AAI60379
21	878.4	30.5	1710	2	AAQ04329
22	876.8	30.4	1717	8	ABA00817
23	860.6	29.9	2385	2	AAV62460
24	665.4	23.1	110000	2	AAT42063_05
25	653.6	22.7	1698	8	ACA45232_05
26	650.4	22.6	2341	8	ACA45080
27	648.2	22.5	1710	10	ADF01572
28	636	22.1	2439	8	ACA46520
29	624	21.7	3919	13	ADT05424
30	624	21.7	85814	13	ADT05644
31	617.2	21.4	2400	2	AAV19002
32	617.2	21.4	4768	3	AAZ92097
33	612.8	21.3	1878	11	ACH96250
34	612.2	21.3	1701	8	ACA35953
35	612.2	21.3	1719	8	ACA34156
36	604.6	21.0	1704	8	ACA25712
37	599	20.8	1719	8	ACA53973
38	598.4	20.8	1704	8	ACA23584
39	586.4	20.4	1698	8	ACA20910
40	585.6	20.3	1704	8	ACA26184
41	582	20.2	1701	8	ACA44198
42	580.2	20.1	110000	10	ACF67367_09
43	580.2	20.1	110000	10	ACF65384_3
44	580	20.1	1716	10	ACF68224
45	573.4	19.9	1704	9	ADA30499

ALIGNMENTS

RESULT 1

ADJ58237

ID ADJ58237 standard; DNA; 2883 BP.

XX AC ADJ58237;

XX DT 06-MAY-2004 (first entry)

XX DE Urease subunit polypeptide complex encoding sequence.

XX KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.

XX OS Helicobacter felis.

XX FH Key Location/Qualifiers

XX FT CDS 206..886

XX FT /*tag= a

XX FT /product= "urease polypeptide complex"

XX FT CDS 897..2603

XX FT /*tag= b

XX FT /product= "urease polypeptide complex"

XX PN EP1176192-A2.

XX PD 30-JAN-2002.

XX PF 11-JUL-2001; 2001EP-00202666.

XX PR 17-JUL-2000; 2000EP-00202565.

XX PA (ALKU) AKZO NOBEL NV.

XX PI Kustérs JG, Cattoli G;

XX DR WPI; 2002-124384/17.

XX DR P-PSDB; ADJ58238, ADJ58239.

XX PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in the diagnosis of Helicobacter felis infections and in the preparation of vaccines.

PS	Claim 1; SEQ ID NO 1; 76pp; English.	
XX	The present invention relates to a novel <i>Helicobacter felis</i> urease X and	
CC	Y subunit polypeptides and immunogenic fragments. The polypeptides are	
CC	used in the manufacture of vaccines against <i>Helicobacter felis</i> infections	
CC	and in diagnostic tests to detect antibodies against <i>Helicobacter felis</i> .	
CC	<i>Helicobacter felis</i> is difficult to grow so it is more convenient to use	
CC	the expression products of the genes encoding the urease X and Y subunits	
CC	in the manufacture of vaccines. The present sequence represents an urease	
CC	subunit polypeptide complex encoding sequence of the invention.	
XX	Sequence 2883 BP; 871 A; 642 C; 666 G; 698 T; 0 U; 6 Other;	
SEQ	Query Match	
	Best Local Similarity 100.0%; Score 2880.6; DB 7; Length 2883;	
	Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 RGGAGATTTCCARCCTTCAAGCACATATTGATCTCTGTGTGGGTGTAATTCR	60
DB	1 RGGAGATTTTCCARCCTTCAAGCACATATTGATCTCTGTGTGGGTGTAATTCR	60
QY	61 ACTTGTTAATCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA	120
DB	61 ACTTGTTAATCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA	120
QY	121 TTACTTATTAATAAAGTTAATAAAGTTAATAAAGTTAATAAAGTTAATAAAGTT	180
DB	121 TTACTTATTAATAAAGTTAATAAAGTTAATAAAGTTAATAAAGTTAATAAAGTT	180
QY	181 TAAATTTAACAAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	240
DB	181 TAAATTTAACAAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	240
QY	241 GTTATATTATCGCGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	300
DB	241 GTTATATTATCGCGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	300
QY	301 ACCGGAAGCCATTGCTTACATTTAGTGGCCCATATTAATGGAAGAGTAAAGTAAAGT	360
DB	301 ACCGGAAGCCATTGCTTACATTTAGTGGCCCATATTAATGGAAGAGTAAAGTAAAGT	360
QY	361 AACCGTGGCCAGCTTATGGAAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	420
DB	361 AACCGTGGCCAGCTTATGGAAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	420
QY	421 CGGGTGGGTAAATGGTTCCGATCTAGGTGTAAGCAGTAAAGTAAAGTAAAGTAAAGT	480
DB	421 CGGGTGGGTAAATGGTTCCGATCTAGGTGTAAGCAGTAAAGTAAAGTAAAGTAAAGT	480
QY	481 ACTTGTAACTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAGCGGCGAAGTGA	540
DB	481 ACTTGTAACTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAGCGGCGAAGTGA	540
QY	541 ATTTGGTTGCGATAAAGCATCGAGTCAATGAGGCAAGTAAAGTAAAGTAAAGTAAAGT	600
DB	541 ATTTGGTTGCGATAAAGCATCGAGTCAATGAGGCAAGTAAAGTAAAGTAAAGTAAAGT	600
QY	601 TACTTAAGAGGCTTAATCTTGCATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAA	660
DB	601 TACTTAAGAGGCTTAATCTTGCATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAA	660
QY	661 CAAGGCCTAAATTCGATCGTGAAGAGCCATGAGCAAGCCCTAGATATTCCTCTGG	720
DB	661 CAAGGCCTAAATTCGATCGTGAAGAGCCATGAGCAAGCCCTAGATATTCCTCTGG	720
QY	721 CAACACGCTACGATTTGGGCGAGGCAAAACCCGCAAGTGAAGTAAAGTAAAGTAAAGT	780
DB	721 CAACACGCTACGATTTGGGCGAGGCAAAACCCGCAAGTGAAGTAAAGTAAAGTAAAGT	780
QY	781 CAGTAAAAAGTGAATGCGATGAAACCGGCTTGTGAATTAACATCGCGGATGAACGCAATA	840
DB	781 CAGTAAAAAGTGAATGCGATGAAACCGGCTTGTGAATTAACATCGCGGATGAACGCAATA	840

QY	841 ACATAAAGCGCTTGACAAGGCGAAATCTCAGCGATTTATCAAGTAAGGAGACTCCCATGA	900
DB	841 ACATAAAGCGCTTGACAAGGCGAAATCTCAGCGATTTATCAAGTAAGGAGACTCCCATGA	900
QY	901 AAATGAAAAACAGAAATATGTAAATATCCTACGACCCACCAAGGCGATAAAGTGGCT	960
DB	901 AAATGAAAAACAGAAATATGTAAATATCCTACGACCCACCAAGGCGATAAAGTGGCT	960
QY	961 TAGGAGATACCGATCTTTGGGCGAAGTAGAATACGACTATATACCACTATGCGGAAGAAC	1020
DB	961 TAGGAGATACCGATCTTTGGGCGAAGTAGAATACGACTATATACCACTATGCGGAAGAAC	1020
QY	1021 TTAATTTGGGCGCGGTAAACTATCCGTGAGGATGGTTCAGAGCAATAGCCCTGATG	1080
DB	1021 TTAATTTGGGCGCGGTAAACTATCCGTGAGGATGGTTCAGAGCAATAGCCCTGATG	1080
QY	1081 AAAACACCTCAGATTTAGTCATCACTAACCGCATGATTATCGACTACACCGGATTTACA	1140
DB	1081 AAAACACCTCAGATTTAGTCATCACTAACCGCATGATTATCGACTACACCGGATTTACA	1140
QY	1141 AAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGGCAAGAAC	1200
DB	1141 AAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGGCAAGAAC	1200
QY	1201 ACATCAAGATCGCGTAAGCCCTCATATGCTGCTGGGTGGGCAAGAGCACTAGCAG	1260
DB	1201 ACATCAAGATCGCGTAAGCCCTCATATGCTGCTGGGTGGGCAAGAGCACTAGCAG	1260
QY	1261 GGGAGGTATGATTATACCGCTGGGGAATCGATTACACACCCCTCTCTTTCTCCAC	1320
DB	1261 GGGAGGTATGATTATACCGCTGGGGAATCGATTACACACCCCTCTCTTTCTCCAC	1320
QY	1321 AACAAATTCCTCCTAGCCCTTAGCCAATGGCGTTTAAACCATTTGGAGGCGGCAAGTCT	1380
DB	1321 AACAAATTCCTCCTAGCCCTTAGCCAATGGCGTTTAAACCATTTGGAGGCGGCAAGTCT	1380
QY	1381 CTGTAGATGGCAAGATGCGACTATCATCTCCGGGCAAAATGGAACTTGCACCGCATGT	1440
DB	1381 CTGTAGATGGCAAGATGCGACTATCATCTCCGGGCAAAATGGAACTTGCACCGCATGT	1440
QY	1441 TCGCGGACAGCAAGATTTCTATGAATGGGCTTTTGGGCAAGGCAATAGCTCTA	1500
DB	1441 TCGCGGACAGCAAGATTTCTATGAATGGGCTTTTGGGCAAGGCAATAGCTCTA	1500
QY	1501 GCAAAAACAACTGTAGAACAAAGTAGAAGCGGCGCGATTTGTTTAAATTTGCATGAAG	1560
DB	1501 GCAAAAACAACTGTAGAACAAAGTAGAAGCGGCGCGATTTGTTTAAATTTGCATGAAG	1560
QY	1561 ACTGGGCGCAACAACCAAGTGGCATCGATCACTGCTTGAGCGTGGCAGATGAATACGATG	1620
DB	1561 ACTGGGCGCAACAACCAAGTGGCATCGATCACTGCTTGAGCGTGGCAGATGAATACGATG	1620
QY	1621 TGCAGTTTGTATCCACACCGGATACAGTCAATGAGGAGGTTATGTAGATGACACCCCTAA	1680
DB	1621 TGCAGTTTGTATCCACACCGGATACAGTCAATGAGGAGGTTATGTAGATGACACCCCTAA	1680
QY	1681 ATGCAATGAACGGGCGCGCATCCATGCTTACCACTTTGAGGAGGCGGTTGGAGGACACT	1740
DB	1681 ATGCAATGAACGGGCGCGCATCCATGCTTACCACTTTGAGGAGGCGGTTGGAGGACACT	1740
QY	1741 CACCTGATTTATCAACATGGCAGGCGAGTCAATATTTCTACCCCTCTCCACCAACCCCTAA	1800
DB	1741 CACCTGATTTATCAACATGGCAGGCGAGTCAATATTTCTACCCCTCTCCACCAACCCCTAA	1800
QY	1801 CTAATTCCTATACCATTAATACGGTTGCAAGACATTTAGACATGCTCATGACATGCCACC	1860
DB	1801 CTAATTCCTATACCATTAATACGGTTGCAAGACATTTAGACATGCTCATGACATGCCACC	1860
QY	1861 ACCTAGACAAACGCAATCCCGGAGGATTTCAATTTTCTCAAGCGGTATCCGCCCGGCT	1920
DB	1861 ACCTAGACAAACGCAATCCCGGAGGATTTCAATTTTCTCAAGCGGTATCCGCCCGGCT	1920
QY	1921 CTATCGCGCTGAAGATGTGCTCATGATATGGGTGTGATCGCGATGACAAGCTCGGATT	1980

```
Db 1921 CTATCGCGCTGAAGATGTGCTCATGATATGGGTGTGATCGGATGCAAGCTCGGATT 1980
Qy 1981 CGCAAGCAATGGCGGTGAGCGGAGAGTGTCTCGAACTTGGCAGACTCGGATAAGA 2040
Db 1981 CGCAAGCAATGGCGGTGAGCGGAGAGTGTCTCGAACTTGGCAGACTCGGATAAGA 2040
Qy 2041 ATAAAAAGATTTGGTAACTTCTGAGAGTGCAGAAAGATAAGCATTAATTCGCAATTA 2100
Db 2041 ATAAAAAGATTTGGTAACTTCTGAGAGTGCAGAAAGATAAGCATTAATTCGCAATTA 2100
Qy 2101 AGCCTACATCTCAAAATACACTATCAACCCGCTTTGACCCAGCGGTGAGCAGATTA 2160
Db 2101 AGCCTACATCTCAAAATACACTATCAACCCGCTTTGACCCAGCGGTGAGCAGATTA 2160
Qy 2161 TCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTGGTGTGGAATCCTCGCTTTTTTG 2220
Db 2161 TCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTGGTGTGGAATCCTCGCTTTTTTG 2220
Qy 2221 GCGTAAACCCAAATCGTGATCAAGGCGGTATGGTGTCTCTGAAATGGGCGATT 2280
Db 2221 GCGTAAACCCAAATCGTGATCAAGGCGGTATGGTGTCTCTGAAATGGGCGATT 2280
Qy 2281 CTAACGCGTCTGTGCCACTCCCAACCGGTTTATTACCGGAAATGTTGGGCATCAG 2340
Db 2281 CTAACGCGTCTGTGCCACTCCCAACCGGTTTATTACCGGAAATGTTGGGCATCAG 2340
Qy 2341 GCAAGGCGAAATTTGACACAGCATCACTTTGTTTCCAAAGTCGCTATGAAATGGCG 2400
Db 2341 GCAAGGCGAAATTTGACACAGCATCACTTTGTTTCCAAAGTCGCTATGAAATGGCG 2400
Qy 2401 TGAAGAAAAGCTGGCTTAGAGCGCAAGTCTACCGGTCAAAACTGCGGTAACTCA 2460
Db 2401 TGAAGAAAAGCTGGCTTAGAGCGCAAGTCTACCGGTCAAAACTGCGGTAACTCA 2460
Qy 2461 CCAAGAAAGACTTCAAGTTCAAAGCAAAACGGCAAAATCACCGTCGATCCGAAACCT 2520
Db 2461 CCAAGAAAGACTTCAAGTTCAAAGCAAAACGGCAAAATCACCGTCGATCCGAAACCT 2520
Qy 2521 TCAGAGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTCTCTAG 2580
Db 2521 TCAGAGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTCTCTAG 2580
Qy 2581 CCAGCGCTACACTTCTCTAGGCAAACTGCGCCCTTTGGGGCAGGTTATTTAGGAA 2640
Db 2581 CCAGCGCTACACTTCTCTAGGCAAACTGCGCCCTTTGGGGCAGGTTATTTAGGAA 2640
Qy 2641 TCTTCATCAAAACGACCTGCAATCGGTCCTTGGTGTGCGATCGTGTGCTTTAAACCAAC 2700
Db 2641 TCTTCATCAAAACGACCTGCAATCGGTCCTTGGTGTGCGATCGTGTGCTTTAAACCAAC 2700
Qy 2701 TTTTCTCTTTAAGCAATCGCAATTTTAAATTAATTTAATTTAATTAATTAATTAAT 2760
Db 2701 TTTTCTCTTTAAGCAATCGCAATTTTAAATTAATTTAATTTAATTAATTAATTAAT 2760
Qy 2761 TATGCCCCCTCATTTTAAAGGAGAAATATGCGTAGGTCTTTGGTATTTGCTATGCGGG 2820
Db 2761 TATGCCCCCTCATTTTAAAGGAGAAATATGCGTAGGTCTTTGGTATTTGCTATGCGGG 2820
Qy 2821 TTGTTTGGTGTGCGGCAAGGATTTGAAACCCATCGCCTCAAAAAGTAGAAGCCAC 2880
Db 2821 TTGTTTGGTGTGCGGCAAGGATTTGAAACCCATCGCCTCAAAAAGTAGAAGCCAC 2880
Qy 2881 AGG 2883
Db 2881 AGG 2883
```

RESULT 2

ADJ58249

ID ADJ58249 standard; DNA; 2452 BP.

XX

AC ADJ58249;

```
XX 06-MAY-2004 (first entry)
DT
XX
DE UreaseXY subunit encoding sequence #4.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
OS Helicobacter felis.
XX
PH Key Location/Qualifiers
CDS 48..728
FT /*tag= a
FT /product= "urease protein"
FT 739..2445
FT /*tag= b
FT /product= "urease protein"
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
XX WPI; 2002-124384/17.
DR P-PSDB; ADJ58250, ADJ58251.
XX
XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
the diagnosis of Helicobacter felis infections and in the preparation of
vaccines.
XX
PS Disclosure; SEQ ID NO 13; 76pp; English.
XX
XX The present invention relates to a novel Helicobacter felis urease X and
Y subunit polypeptides and immunogenic fragments. The polypeptides are
used in the manufacture of vaccines against Helicobacter felis infections
and in diagnostic tests to detect antibodies against Helicobacter felis.
XX
XX Helicobacter felis is difficult to grow so it is more convenient to use
the expression products of the genes encoding the urease X and Y subunits
in the manufacture of vaccines. The present sequence represents an
ureaseXY subunit of the invention.
XX
SQ Sequence 2452 BP; 746 A; 568 C; 590 G; 548 T; 0 U; 0 Other;
Query Match 77.3%; Score 2228; DB 7; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 159 AGGACTATAATCCCATTCCTTTAAATTTAAACAAGAGTAAATAGTGAACTCACAC 218
Db 1 AGGACTATAATCCCATTCCTTTAAATTTAAACAAGAGTAAATAGTGAACTCACAC 60
Qy 219 CCAAGAGCAAGAAAAGTTCTTGTATATATATGCGGCGAAGTGGCTAGAAAGCGCAAG 278
Db 61 CCAAGAGCAAGAAAAGTTCTTGTATATATGCGGCGAAGTGGCTAGAAAGCGCAAG 120
Qy 279 CAGAGGCTTAAAGCTCAACCAACCCGAAAGCATTGCTTACATTAGTCCCATATATGCG 338
Db 121 CAGAGGCTTAAAGCTCAACCAACCCGAAAGCATTGCTTACATTAGTCCCATATATGCG 180
Qy 339 ACGAAGCGCGCGTGGAAAAGAAAACCGTTGCGCCAGCTTATGGAAGAGTGCATGCACTTTT 398
Db 181 ACGAAGCGCGTGGGCAAGAAAACCGTTGCGGAACTTATGGAAGAGTGTATGCACTTTT 240
Qy 399 TGAAGAAAAGATGAAGTAATGCCCGGGTGGGTATATATGTTCCCGATCTAGTGTAGAAAG 458
Db 241 TGAAGAAAAGACGAGGTGATGTCGCCGGGTGGGGAATATGTTCCCTGATTGGGCGTGAAG 300
```

QY 459 CCACCTTTCCTGATGGTACGAACTTGTAACCTGTGAATTTGGCCCATCGAACACAGATGACC 518
Db 301 CCACCTTCCCAGATGGCCACCAAACTCGTAACCTGTGAATTTGGCCCATCGAACCTGTATGAC 360
QY 519 ACTTCAAAAGCGGCGAAGTGAAATTTGGTTGGATAAAGACATCGAGCTCAATGACGCA 578
Db 361 ACTTTAAGCGGCTGAAGTGAAATTTGGCTGTGTATTAAGACATTTGAACTCAACGCAAGTA 420
QY 579 AAGAAGTAACCGAACTTGAGGTTACTAATGAAGGCGCTAAATTCCTTGATGTGGTAGCC 638
Db 421 AGGAAGTTACCGAACTAGAAGTTACTAAGGAAGGACCTAAATTCCTTGATGTGGTAGCC 480
QY 639 ATTTCCACTTCTTTGAAGCTAACAAGGCACTTAAATTCGATCGTGAAAGGCTTAGGCA 698
Db 481 ATTTCCACTTCTTTGAAGCACAACAAAGCATTTGAAATTCGATCGGGAAGGCTTAGGCA 540
QY 699 AACGCTAGATATTCCTCTGCGCAACGCTACGCAATTTGGGCGACGACAAACCCGCAAG 758
Db 541 AACGCTAGATATTCCTCTGCGCAACGCTACGCAATTTGGGCGACGACAAACCCGTAAG 600
QY 759 TGCAGTTGATTCCTCTGCGTGCAGTAAAAAGTGAATTTGGCATGAACGGGCTTTGTGAATA 818
Db 601 TGCAGTTAATCCCTCTTGGCGTAGTAAAAAGTGAATTTGGCATGAACGGGCTTTGTGAATA 660
QY 819 ACATGCGCGATGAAGCCCATAAACATAAAGCGCTTGACAAAGCGGAAATCTCACGGATTTA 878
Db 661 ATAATTGCGGAGCAAGCCCATAAACAAAGCGCTAGACAAAGCAAAATCTCACGGATTTA 720
QY 879 TCAAGTAGGAGACTCCCATGAATGAATAAACAAGAAATGTAAATACCTACCGACCC 938
Db 721 TCAAGTAGGAGACTCCCATGAATGAATAAACAAGAAATGTAAATACCTACCGACCC 780
QY 939 ACCAAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGAAGTAGAAACATGAC 998
Db 781 ACCAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGAAGTAGAAACATGAC 840
QY 999 TATACCACTATGCGGAAGAACTTAAATTTGGCGGGGTAAACATCTCCGTGAGGGTATG 1058
Db 841 TATACCACTATGCGGAAGAACTCAAATTCGCTGCAAGGTAAACATCTCCGTGAGGGTATG 900
QY 1059 GGTACAGCATAGCCCTGATGAAACACCCCTAGATTTAGTTCATCTAACCGCATGATT 1118
Db 901 GGTACAGCATAGCCCAAGTGAACACCTTAGATTTAGTTCATCTAACCAACGCGATGATT 960
QY 1119 ATCGACTACCGGGATTTACAAAGCGACATTTGGGATTTAAACCGCAAAATCCATGCG 1178
Db 961 ATTGACTACCGGGATTTACAAAGCGACATTTGGCATTTAAAAATGGCAAAATCCATGCG 1020
QY 1179 ATTGCAAGGCGAAGAAACAAAGACATGCAAGATGCGGTAAGCCCTCATATGTCGTGGGT 1238
Db 1021 ATTGCAAGGCGAAGAAACAAAGACATGCAAGATGCGGTAAGCCCTCATATGTCGTGGGT 1080
QY 1239 GTGGGCAAGAGCACTAGAGGGGAAGGTATGATTTACCGCTGGGGGAATCGAATCA 1298
Db 1081 GTGGGCAAGAGCACTAGAGGGGAAGGTATGATTTACCGCTGGGGGAATCGAATCA 1140
QY 1299 CACACCACTTCTTCTTCCACAAATTCCTACCGCTCTAGCCCAATGGGGTTACAACC 1358
Db 1141 CACACCACTTCTTCTTCCACAAATTCCTACCGCTCTAGCCCAATGGGGTTACAACA 1200
QY 1359 ATGTTTGGAGGCGGACAGGCTCTGTAGATGGCAGCAATGCGACTACTATCACTCCGGGC 1418
Db 1201 ATGTTTGGAGGCGGACAGGCCCCGTAGATGGCAGCAATGCGACTACTATCACTCCGGGC 1260
QY 1419 AAATGGAACTTGACCCGATGTTGGCGGAGCAGAAAGATTTCTATGAATGTGGGCTTT 1478
Db 1261 AAATGGAACTTGACCCGATGTTGGCGGAGCAGAAAGATTTCTATGAATGTGGGCTTT 1320
QY 1479 TTGGGCAAGGCAATAGCTCTAGCAAAACAACTTTGTAGAACAGTAGAGCGGGCGG 1538
Db 1321 TTGGGCAAGGCAATAGCTCTAGTAAAAACAACTTTGTAGAACAGTAGAGCGGGCGG 1380
QY 1539 ATTGGTTTTAAATTGCATGAAGAAGCTGGGGCAACACCAAGATGCGATCGATCACTGCTTG 1598

Db 1381 ATTGGTTTTAAATTGCATGAAGAAGCTGGGGCAACACTCCAAAGTCCGATCGATCACTGCTTG 1440
QY 1599 AGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACGATACAGTCAATGAGGCA 1658
Db 1441 AGCGTGGCAGATGAATACGATGTGCAAGTTTGTATACACCGATACGATCAATGAGGCA 1500
QY 1659 GGTATGTAGATGAACACCTAAATGCAATGAACGGCGGCCCATCCATGCTTACCACTT 1718
Db 1501 GGTATGTAGATGAACACCTAAATGCAATGAACGGCGGCCCATCCATGCTTACCACTT 1560
QY 1719 GAGGAGCGGTTGGAGGACACTCACTGATGTATACCATTAACGAGCGAGCTCAATATT 1778
Db 1561 GAGGAGCGGTTGGAGGACACTCACTGATGTATACCATTAACGAGCGAGCTCAATATT 1620
QY 1779 CTACCTCTCTCCACACCCCACTATTCCTATACCATTAACGAGTTGAGAACTTAA 1838
Db 1621 CTACCTCTCTCCACACCCCACTATTCCTATACCATTAACGAGTTGAGAACTTAA 1680
QY 1839 GACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGGATTTACAAATTTCT 1898
Db 1681 GACATGCTTATGACCTGCGCACCATCTAGATAAACGATCCGCGAGGATCTCCAATTTCT 1740
QY 1899 CHAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGTCTCCATGATATGGGTG 1958
Db 1741 CHAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGTCTCCATGATATGGGTG 1800
QY 1959 ATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGTGCAAGGCAAGTGTCTCTCGA 2018
Db 1801 ATCGCGATGACAAGTTCGATTCGCAAGCAATGGGCGCGCTGGGGAAGTGTCTCTAGA 1860
QY 2019 ACTTGGCAGACTCGGATGAAGATAAAGAAATTTGGTAAGCTTCTCTGAAGATGCAAA 2078
Db 1861 ACTTGGCAAACTCGACAGCAAGAAATAAAGAAATTTGGTAAGCTTCTCTGAAGATGCA 1920
QY 2079 GATACGATAATTTCCGATTAAGCGCTACATCTCAAAATACATATCAACCCCGCTTTG 2138
Db 1921 GATAATGACAACTTCGCGATCAACGCTATATCTCCAAATACACCATAATTCGCGCTTG 1980
QY 2139 ACCCAGCGGTGAGGAGTATATCGCTCTGTGGAAGGCGCAAGATCCCGACTTGGTG 2198
Db 1981 ACCCATGGGTGAGGAGTATATCGCTCTGTGGAAGGCGCAAGATCCCGACTTGGTG 2040
QY 2199 GTTGGAAATCTCTGCTTTTGGCGTAAACCCAAATCGTATCAAAAGCGGTATGGTG 2258
Db 2041 GTTGGAAATCTCTGCTTTTGGCGTAAACCCAAATCGTATCAAAAGCGGTATGGTG 2100
QY 2259 GTCTTCTCTGAAATGGCGATTTAAACGCGCTCTGTGCCCATCTCCCAACCGGTTTATTAC 2318
Db 2101 GTGTTCTCTGAAATGGCGATTTCTAATGCTCTGTGCCACTCTCTCAGCGGTTATTAC 2160
QY 2319 CGCGAAATGTTTGGGCATCACGGCAAGGCAAAATTTGACACGAGCATCACTTTTGTTC 2378
Db 2161 CGCGAAATGTTTGGGCATCACGGCAAGGCAAAATTTGACACGAGCATCACTTTTGTTC 2220
QY 2379 AAAAGTCGCTATGAAATGGCGTGAAGAAAGCTTGGGCTTAGAGCGCCAAAGTTCTACCG 2438
Db 2221 AAAAGTCGCTATGAAATGGGTGTGAAGAAAGCTTAGGTTTAGAGCGCAAGTGTCTCCC 2280
QY 2439 GTCAAAAATGCGCGTAAACATCAACGAAGAAAGCTTCAAGTTCAACGAACAAACGGCAAAA 2498
Db 2281 GTCAAAAATGCGCGTAAACATCAACGAAGAAAGCTTCAAGTTCAACGAACAAACGGCAAAA 2340
QY 2499 ATCACCGGTGATCCGAAACCTTCAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2558
Db 2341 ATCACCGGTGATCCGAAACCTTCAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2400
QY 2559 CCACCTGCAAGTGCCTCTAGCCAGGCTACACTTTTCTTAGGCACAAAT 2610
Db 2401 CCACCTGCAAGTGCCTCTAGCCCAACGCTAACCTTTCTTAGGCATAAT 2452

```
ADJ58240
ID ADJ58240 standard; DNA; 2405 BP.
XX AC ADJ58240;
XX DT 06-MAY-2004 (first entry)
XX DE UreaseXY subunit encoding sequence #1.
XX KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX OS Helicobacter felis.
XX FH Key Location/Qualifiers
XX CDS 1..681
XX FT /tag= a
XX FT /product= "urease protein"
XX FT 692..2398
XX FT /tag= b
XX FT /product= "urease protein"
XX PN EP1176192-A2.
XX PD 30-JAN-2002.
XX PF 11-JUL-2001; 2001EP-00202666.
XX PR 17-JUL-2000; 2000EP-00202565.
XX PA (ALKU ) AKZO NOBEL NV.
XX PI Kusters JG, Cattoli G;
XX WPI: 2002-124384/17.
XX P-PSDB; ADJ58241, ADJ58242.
XX PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
XX the diagnosis of Helicobacter felis infections and in the preparation of
XX vaccines.
XX PS Disclosure; SEQ ID NO 4; 76pp; English.
XX CC The present invention relates to a novel Helicobacter felis urease X and
XX Y subunit polypeptides and immunogenic fragments. The polypeptides are
XX used in the manufacture of vaccines against Helicobacter felis infections
XX and in diagnostic tests to detect antibodies against Helicobacter felis.
XX Helicobacter felis is difficult to grow so it is more convenient to use
XX the expression products of the genes encoding the urease X and Y subunits
XX in the manufacture of vaccines. The present sequence represents an
XX ureaseXY subunit of the invention.
XX SQ Sequence 2405 BP; 715 A; 571 C; 588 G; 531 T; 0 U; 0 Other;

Query Match 74.9%; Score 2158.6; DB 7; Length 2405;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 206 GTGAACTCACCACCAAGAGCGGCTTAAAGCTCAACCAACCCGAAGCATTGCTTACATTAGT 265
DB 1 GTGAACTCACCACCAAGAGCGGCTTAAAGCTCAACCAACCCGAAGCATTGCTTACATTAGT 60
QY 266 AGAAAGCGCAAGAGAGGCGCTTAAAGCTCAACCAACCCGAAGCATTGCTTACATTAGT 325
DB 61 AGAAAGCGCAAGAGAGGCGCTTAAAGCTCAACCAACCCGAAGCATTGCTTACATTAGT 120
QY 326 GCCCATATTGACGAGCGCGCGTGGAAAAAACCCTTGCCAGCTTATGGAAGAG 385
DB 121 GCCCATATTGACGAGCGCGCGTGGAAAAAACCCTTGCTGAACCTTATGGAAGAA 180
QY 386 TGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCGGGTGGTATATGTTCCCGAT 445
DB 181 TGTATGCACCTTTTGAAGAAAGATGAAGTATGCCGGGTGGTATATGTTCCCTGAT 240
```

```
Db 1321 AGAAGCGGCGCGATGGCTTTAAATGCAAGACTGGGGCAACACCAAGTGGCAT 1380
Qy 1586 CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTGCTTGAGCGTAGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1440
Qy 1646 AGTCAATGAGCGAGTTATGTAGATGACACCTAAATCAATGAACGGCGCGCATCCA 1705
Db 1441 GGTCAATGAGCGAGTTATGTAGATGACACCTAAATCGATGAACGGCGCGCATCCA 1500
Qy 1706 TGCCTACCAATGAGCGAGCGGTGGAGGACACTCACTGATGTATATCAATGGCAGG 1765
Db 1501 TGCCTACCAATGAGCGAGCGGTGGAGGACACTCACTGATGTATATCAATGGCAGG 1560
Qy 1766 CGAGCTCAATATCTACCTCTCCACACCCCACTATTCCTATACCATTAATACCGT 1825
Db 1561 CGAGCTCAATATCTACCTCTCCACACCCCACTATTCCTATACCATTAATACCGT 1620
Qy 1826 TGCAGAACATTTAGCATGCTCATGACATGCCACCTAGACAAACGCATCCGCGAGGA 1885
Db 1621 TGCAGAACATTTAGCATGCTCATGACATGCCACCTAGACAAACGCATCCGCGAGGA 1680
Qy 1886 TTTCAATTTTCTCAAGCCGATATCCGCGCGCTCTATCGCGCTGAAGATGCTCCA 1945
Db 1681 TCTCCAGTTTCCCAAGCCGATATCCGCGCGCTCTATCGCGCTGAAGATGCTCCA 1740
Qy 1946 TGATATGGTGTGATCGGATGACAAGCTCGGATTCGAAGCAATGGGGCGTGACAGCGA 2005
Db 1741 TGATATGGTGTGATCGGATGACAAGCTCGGATTCGAAGCAATGGGGCGTGAGGA 1800
Qy 2006 AGTGATTCCTCGAACTTGGCAGACTGCGGATGAAGATAAAGAAATTTGGTAAGCTTCC 2065
Db 1801 AGTGATTCCTAGAACTTGGCAAACTGCAGACAAAGAAATTAAGAAATTTGGTAAGCTTCC 1860
Qy 2066 TGAAGATGGCAAGATTAACGATTAATTTCCGATTAAGCGCTACATCTCCAATACACTAT 2125
Db 1861 TGAAGATGGTGCAGATTAACGATTAATTTCCGATTAAGCGCTACATCTCCAATACACTAT 1920
Qy 2126 CAACCCCGCTTGAACCGCGTGGAGGATATATCGGCTCTGGAAGAGGGCAAGAT 2185
Db 1921 TAATCCCGCTTGAACCGCGTGGAGGATATATCGGCTCTGGAAGAGGGCAAGAT 1980
Qy 2186 CGCGACTTGGTGTGATGGAATCTGCTTTTGGCGTAAACCCAAATCGTGATCAA 2245
Db 1981 CGCGACTTGGTGTGATGGAATCTGCTTTTGGCGTAAACCCAAATCGTGATCAA 2040
Qy 2246 AGCGGATATGGTGTCTCTGAAATGGGCGATTTCTAAGCGCTGTGCGCCACTCCCA 2305
Db 2041 AGCGGATATGGTGTCTCTGAAATGGGCGATTTCTAAGCGCTGTGCGCCACTCCCA 2100
Qy 2306 ACCGCTTTATACCGGAAATTTTGGGCATCACGCAAGCGAAATTTGACACCGCAT 2365
Db 2101 GCCGCTTTATACCGGAAATTTTGGGCATCACGCAAGCGAAATTTGACACCGCAT 2160
Qy 2366 CACTTTTGTTCCTCAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTTGTTCCTCAAGTCGCTATGAAATGGCGTGAAGAAAGCTTAGAGCG 2220
Qy 2426 CCAAGTTCTACCGGTCAAAAATCGCGTAAACATCAACGAAGAAGACTTCAAGTTCAACGA 2485
Db 2221 CAAGGTGTACCGGTGA AAAATCGCGCAACATCACTAAGAAGACTTCAAAATCAACAA 2280
Qy 2486 CAAACGGCAAAATCACCGTCGATCCGAAACCTTCGAGTCTTTGTAGATGGCAACT 2545
Db 2281 CAAGCGCGCATATCACTGTGATCTTAAACCTTCGAGGCTTTGTAGATGGCAACT 2340
Qy 2546 CTGCACTCTAAACCCACTCGAGTGCCTTAGCCAGCGCTACACTTTCTCTAGGC 2605
Db 2341 CTGCACTCTAAACCCCGCTCTGAAGTGCTCTTAGCCCAACGCTACACTTTCTTAGGC 2400
Qy 2606 ACAAT 2610
|||||
```

```
Db 2401 ACAAT 2405

RESULT 4
ADJ58246
ID ADJ58246 standard; DNA; 2407 BP.
XX
AC ADJ58246;
XX
DT 06-MAY-2004 (first entry)
XX
UreaseX subunit encoding sequence #3.
XX
immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
Helicobacter felis.
XX
Key Location/Qualifiers
CDS 2..682
FT /*tag= a
FT /product= "urease protein"
FT 693..2399
FT /*tag= b
FT /product= "urease protein"
XX
EP1176192-A2.
XX
30-JAN-2002.
XX
11-JUL-2001; 2001EP-00202666.
XX
17-JUL-2000; 2000EP-00202565.
XX
(ALKU ) AKZO NOBEL NV.
XX
Kusters JG, Cattoli G;
XX
WPI; 2002-124384/17.
XX
P-PSDB; ADJ58247, ADJ58248.
XX
Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
the diagnosis of Helicobacter felis infections and in the preparation of
vaccines.
XX
Disclosure; SEQ ID NO 10; 76pp; English.
XX
The present invention relates to a novel Helicobacter felis urease X and
Y subunit polypeptides and immunogenic fragments. The polypeptides are
used in the manufacture of vaccines against Helicobacter felis infections
and in diagnostic tests to detect antibodies against Helicobacter felis.
XX
Helicobacter felis is difficult to grow so it is more convenient to use
the expression products of the genes encoding the urease X and Y subunits
in the manufacture of vaccines. The present sequence represents an
ureaseX subunit of the invention.
XX
SQ Sequence 2407 BP; 720 A; 574 C; 584 G; 528 T; 0 U; 1 Other;

Query Match 74.9%; Score 2158.6; DB 7; Length 2407;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 206 GTGAAACTCACACCCCAAGAGCAAGAAAGTTCTTGTTATATATCGCGCGAAGTGGCT 265
Db 2 GTGAAACTCACACCCCAAGAGCAAGAAAGTTCTTGTTATATATCGCGCGAAGTGGCT 61
Qy 266 AGAAGCGCAAGCAGAGCGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 325
Db 62 AGAAGCGCAAGCAGAGCGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 121
Qy 326 GCCCATATTATGAGCAAGCGCGCTGGAAGAAACCGTTGCCAGCTTATGGAGAG 385
Db 122 GCCCATATTATGAGCAAGCGCGCTGGAAGAAACCGTTGCCAGCTTATGGAGAG 181
```


QY 386 TGCATGCACTTTTGGAAAAAGATGAAGTAAATGCCGGGTGGGTAAATATGTTCCCGAT 445
DB 182 TGTATGCACTTTTGGAAAAAGACAGAGTGATGCCGGGTGGGTAAATATGTTCCCTGAT 241
QY 446 CTAGGTGTAGAGCCACCTTTCTCTGATGGTACGAAACTTGTAACTGTGAAATTTGGCCCATC 505
DB 242 TTAGCGGTGGAAGCTACTTTTCCCGATGGCACCAAACTCGTAAACGTTGAAATTTGGCCCATC 301
QY 506 GAAACAGATGAGCACTTCAAAGCGGGCGAAGTGAATTTGGTTGCGATAAGACATCGAG 565
DB 302 GAAACCGATGAACACTTCAAAGCGGGCGAAGTCAAATTTGGCTGTGATAAAGACATTGAA 361
QY 566 CTCAATGAGCAAGAAAGTAAACCGAACTTGAGGTTACTTAATGAAGGCGCTAAATTCCTTG 625
DB 362 CTCAAGCGAGTGAAGAGTTACCGAAGTTAGAACTTCAAGCAAGGACCTAAATTCCTTG 421
QY 626 CATGTGGGTAGCCATTTTCCACTTTCTTGAAGCTAAAGAGCACTAAATTCGATCGTGAA 685
DB 422 CATGTGGGTAGCCATTTTCCACTTTCTTGAAGCTAAAGAGCACTAAATTCGATCGTGAA 481
QY 686 AAAGCCTATGAGCAAGCGCTAGATATTCCTCTCTGGTGGCAATAAAGTGAATGGCATGAAC 805
DB 482 AAAGCCTATGAGCAAGCGCTAGATATTCCTCTCTGGTGGCAATAAAGTGAATGGCATGAAC 601
QY 806 GGGCTTGTGAATAACATCGCGGATGAACGCGCATAAACATAAAGCGCTTGACAAGGCGAAA 865
DB 602 GGGCTTGTGAATAATTTGCAAGTGAAGCGCATAAACATAAAGCGTTAGAAAAGCAAAA 661
QY 866 TCTCAGCGATTTTCAAGTAAAGGAGCTCCCATGAAATGAAAAACAAAGATATGTAAA 925
DB 662 TCTCAGCGATTTTCAAGTAAAGGAGCTCCCATGAAATGAAAAACAAAGATATGTAAA 721
QY 926 TACCTAGCGACCCACCAAGGCGATAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGA 985
DB 722 TACCTAGCGACCTTACCAAGGCGCAAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGA 781
QY 986 AGTAGAATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGTAAACCTAT 1045
DB 782 AGTAGAATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGTAAACCTAT 841
QY 1046 CCGTAGGGGTATGGGTGAGAGCAATAGCCCTTGATGAAAAACACCCCTAGATTTAGTCACTAC 1105
DB 842 CCGTAGGGCATGGGTGAGAGCAATAGTCCAGATGAAAAACACCCCTAGATTTAGTCACTAC 901
QY 1106 TAAGCGGATGATTATGCACTACACCGGATTTTCAAAGCGGACATTTGGGATTAATAAAGCG 1165
DB 902 CAACCGGATGATTATGCACTACACCGGATTTTCAAAGCGGACATTTGGGATTAATAAAGCG 961
QY 1166 CAATAATCCATGGCATTTGGCAAGGCGAGAAACAAAGCAATGCAAGATGGCGTAAAGCCCTCA 1225
DB 962 CAATAATCCATGGCATTTGGCAAGGCGAGAAACAAAGCAATGCAAGATGGCGTAAAGCCCTCA 1021
QY 1226 TATGGTCTGGGTGGGCAAGAAAGCACTAGCAGGAGGAGGTATGATTTACCGCTGG 1285
DB 1022 TATGGTCTGGGTGGGCAAGAAAGCACTAGCAGGAGGAGGTATGATTTACCGCTGG 1081
QY 1286 GGGAAATCGATTACACACCCACTTTCCTTTCTCAACAAATTCCTACCGCTCTAGCCAA 1345
DB 1082 GGGATCGATTACACACCCACTTTCCTTTCTCAACAAATTCCTACCGCTCTAGCCAA 1141
QY 1346 TGGGTTTACACCAATTTTGGAGGCGGCAAGGTCTCTGTAGATGGCAAGATGGCACTAC 1405
DB 1142 TGGGTTTACACCAATTTTGGAGGCGGCAAGGTCTCTGTAGATGGCAAGATGGCACTAC 1201
QY 1406 TATCACTCCGGGCAAAATGGAATTTGACCGCATGTTGCGGCGAGCAAGAGTATTTCTAT 1465
DB 1202 CATCACTCCGGGCAAAATGGAATTTGACCGCATGTTGCGGCGAGCTGAAGAGTATTTCTAT 1261
QY 1466 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACTTGTAGAACAAAT 1525

DB 1262 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCCAGTAAAAACAACTTGTAGAACAAAT 1321
QY 1526 AGAAGCGGCGCGATTTGGTTTAAATTTGCATGAAGACTGGGGCAACCAAGTGGCAT 1585
DB 1322 AGAAGCGGCGCGATCGGCTTTAAATTTGCATGAAGACTGGGGCAACCTCAAAGTGCAT 1381
QY 1586 CGATCACTGCTTTAGCGGTGGCAGATGAATACGATGTCGAAGTTTGTATCCACACCGATAC 1645
DB 1382 CGATCACTGCTTTAGCGGTAGCAGATGAATACGATGTCGAAGTTTGTATCCACACCGATAC 1441
QY 1646 AGTCAATGAGGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAAAGGGCGGCCATCCA 1705
DB 1442 GGTCAATGAGGCGAGGTTATGTAGATGACACCCCTGAATGCGATGAACGGGGCGGCCATCCA 1501
QY 1706 TGCCTACCAATTTGAGGGAGCGGTGGAGGACACTCACCTGATGTTTATCAACATGGCAGG 1765
DB 1502 TGCCTACCAATTTGAGGGAGCGGTGGAGGACACTCACCTGATGTTTATCACCATGGCAGG 1561
QY 1766 CGAGCTCAATATTTCTACCTCTCTGACACCCCGCTATTTCCCTATACCATTAATACGGT 1825
DB 1562 CGAGCTCAATATTTCTACCTCTCTGACCAACCCCGCTATTTCCCTATACCATTAATACGGT 1621
QY 1826 TGCAGAACACTTTAGACATGCTCATGACATGCCACACCTAGACAAACGATCCGCGAGGA 1885
DB 1622 TGCAGAACACTTTAGACATGCTCATGACCTGACACCTAGATAAACGATCCGCGAGGA 1681
QY 1886 TTTTCAATTTTCTCAAAGCGGTATCGCCCGCGCTCTATCGCGGCTGAAGATGTGTCTCA 1945
DB 1682 TTTTCAATTTTCTCAAAGCGGTATCGCCCGCGCTCTATCGCGGCTGAAGATGTGTCTCA 1741
QY 1946 TGATATGGTGTGATCGCGATGACAAAGCTCGAATTCGAAAGCAATGGGGGTGCAGGCGA 2005
DB 1742 TGATATGGCGGTGATCGCGATGACAAAGCTCGAATTCGAAAGCAATGGGGGTGCAGGCGA 1801
QY 2006 AGTGATTTCTCGAACTTTGGCAGACTCGCGATGAAGTAAAGAAATTTGGTAAAGCTTC 2065
DB 1802 AGTGATTTCTCGAACTTTGGCAGACTCGCGATGAAGTAAAGAAATTTGGTAAAGCTTC 1861
QY 2066 TGAAGATGGCAAGATTAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTAT 2125
DB 1862 TGAAGATGATGAGATTAACGACACTTTCCGTATCAACCGCTACATCTCCAAATACACTAT 1921
QY 2126 CAACCCCGCTTTGACCCACGGGTGAGCGAGTATATCGGCTCTGTGGAAAGAGGCGCAAGAT 2185
DB 1922 TAAACCCCGCTCTAACCCATGGGTAAAGCGAGTATATCGGCTCTGTGGAAAGAGGCGCAAAAT 1981
QY 2186 GCGGCACTTGGTGGTGAATCTCGCTTTTGGCGGTAAACCCCAAAATCGTGAATCAA 2245
DB 1982 CGTGATTTGGTGGTGGAAATCTCGCTTTTGGTGTGAAACCTTAAGATTTGTGATCAA 2041
QY 2246 AGGCGTATGGTGGTCTTCTCTGAAATGGGCGATTTCTAACGCGTCTGTCCCACTCCCCA 2305
DB 2042 AGGCGTATGGTGGTCTTCTCTGAAATGGGCGACTTCAACGCGTCCGTGCTTACACTCA 2101
QY 2306 ACCGCTTTATTAACCGCGAAATGTTTGGGCACTCAACGCGGAGGCGAAATTTGACACCAAGCAT 2365
DB 2102 GCGGCTTTATTAACCGCGAAATGTTTGGGCACTCAACGCGGAGGCGAAATTTGACACCAAGCAT 2161
QY 2366 CACTTTTGTGTTTCCAAAGTGGCTATGAAAAATGGCGTGAAGAAAAAGCTGGGCTTAGAGCG 2425
DB 2162 CACTTTTGTGTTTCCAAAGTGGCTATGAAAAATGGCGTGAAGAAAAAGCTTAGGCTTAGAGCG 2221
QY 2426 CCAAGTTCTACGGGTCAAAAACTGGGTGAACATCAACCAAGAAAGCTTCAAGTTCAACGA 2485
DB 2222 CAAAGTGTACCCCGTGAAAAACTGGCGCAACATCACTAAGAAAGACTTCAAAATTCACAA 2281
QY 2486 CAAAACGGCAAAAACTCACCGCTCGATCCGAAACCTTTGAGGCTTTGTAGATGGCAAACT 2545
DB 2282 CAAAGCGGCGATATCACTGTGATCTTAAACCTTTGAGGCTTTGTAGATGGCAAACT 2341
QY 2546 GTGACCTCTTAAACCCACCTCGCAAGTGGCTCTAGCCCAAGCGCTACATTTTCTTAGGC 2605

Db	2342	CTGCACCTCTAAACCGCCTCTGAAGTGCTCTAGCCAGCGGTACACTTCTTCTTAGGC	2407
Qy	2606	ACAATG 2611	
Db	2402	NCAATG 2407	
RESULT 5			
ADJ58243	ADJ58243 standard; DNA; 2183 BP.		
XX	ADJ58243;		
XX	06-MAY-2004	(first entry)	
DT	UreaseXY subunit encoding sequence #2.		
XX	immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.		
KW	Helicobacter felis.		
OS			
XX	Key	Location/Qualifiers	
FF	CDS	3. .683	
FT		/*tag= a	
FT		/product= "urease protein"	
FT		594. .2181	
FT		/*tag= b	
FT		/product= "urease protein"	
XX	EP1176192-A2.		
XX	30-JAN-2002.		
XX	11-JUL-2001; 2001EP-00202666.		
XX	17-JUL-2000; 2000EP-00202565.		
XX	(ALKU) AKZO NOBEL NV.		
PA	Kusters JG, Cattoli G;		
XX	WPI; 2002-124384/17.		
DR	P-PSDB; ADJ58244, ADJ58245.		
XX	Novel Helicobacter felis urease X and Y subunit polypeptides, useful in		
PPT	the diagnosis of Helicobacter felis infections and in the preparation of		
PT	vaccines.		
XX	Disclosure; SEQ ID NO 7; 76pp; English.		
XX	The present invention relates to a novel Helicobacter felis urease X and		
CC	Y subunit polypeptides and immunogenic fragments. The polypeptides are		
CC	used in the manufacture of vaccines against Helicobacter felis infections		
CC	and in diagnostic tests to detect antibodies against Helicobacter felis.		
CC	Helicobacter felis is difficult to grow so it is more convenient to use		
CC	the expression products of the genes encoding the urease X and Y subunits		
CC	in the manufacture of vaccines. The present sequence represents an		
CC	ureaseXY subunit of the invention.		
XX			
SQ	Sequence 2183 BP; 638 A; 505 C; 550 G; 490 T; 0 U; 0 Other;		
Query Match 67.6%; Score 1948; DB 7; Length 2183;			
Best Local Similarity 93.3%; Pred. No. 0;			
Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;			
Qy	206	GTGAAACTCACACCCAAAGACGAAAGTCTTGTATATATTATGCGGCGAAGTGCT 265	
Db	3	GTGAAACTCACACCCAAAGACGAAAGTCTTGTATATATTATGCGGCGAAGTGCT 62	
Qy	266	AGAAAGCGGAAAGCAGAGGCGTTAAAGCTCAACCAACCGAGCCATTGCTTACATTAGT 325	
Db	63	AGAAAGCGGAAAGCAGAGGCGTTAAAGCTCAATCAACCGAGCCATTGCTTACATTAGT 122	

QY		326	GCCCATATTATGACGAAGCGCGCTGTGGAAAAAACCCTTGCCTCAGCTTATGTGAAGAG	385
Db		123	GCCCATATTATGACGAAGCGCGCTGTGGAAAAAACCCTTGCCTCAGCTTATGTGAAGAA	182
QY		386	TGCATGCACCTTTTGAATAAAGATGAAGTAACTGCCCGGGTGGGTAATATGGTTCCTCGAT	445
Db		183	TGTATGCACTTTTTGAATAAAGATGAGTGATGCCCGGTGTGGGGAATATGGTCCCTGAT	242
QY		446	CTAGGTGTAGAAGCCACTTTCTCGATGTGTACGAACCTTTGTAACTGTGAATTTGGCCCCATC	505
Db		243	TTGGCGGTAGAAGCCACTTTCCC CGATGGCACCAAACCTCGTAACCGTGAATTTGGCCCCATT	302
QY		506	GAACCHAGTAGACACTTCAAGCGGCGGAAGTGAAAATTTGGTTCGATATAAGACATCGAG	565
Db		303	GAACTTGATGAACACTTTTAAAGCGGCTGAAGTGAATTTGGCTGTGATAAAGACATTTGAG	362
QY		566	CTCAATCGAGGCAAGAGTAAACCGAACCTTGAGGTTTACTAAATGAAGGGCCCTAAATCCTTG	625
Db		363	CTCAAGCTGGGTAAAGGAAGTTTCCGAGCTTTGAAGTTACCAACGAGGACCTAAATCCTTG	422
QY		626	CATGTGGGTAGCCATTTCACATCTCTTTTGAAGCTTAACAAGSCACTAAATTCGATCGTGAA	685
Db		423	CATGTGGGTAGCCATTTCACATCTCTTTTGAAGCAACAAGSCATTGAAATTCGATCGGAA	482
QY		686	AAGCCTATGGCAAAAGCCTAGATATTCCTCTGGCAACAACGCTACGCATTTGGGGCAGGA	745
Db		483	AAGCCTATGGCAAAAGCCTAGATATTCCTCTGGCAACAACGCTACGCATTTGGGGCAGGA	542
QY		746	CAAAACCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAGTGAATTCGATCGAAC	805
Db		543	CAAAACCGTAAAGTCAGTTAAATTCCTCTTGGCGGTAGTAAAAAGTGAATTCGATCGAAC	602
QY		806	GGCTTGTGAATAACATCGCGATGAAGCCATAAACAATAAAGCGCTTGA CAAGGCGAAA	865
Db		603	GGCTTGTGAATAATTCGCGAGCAAGCCATAAACAATAAAGCACTAGACAAGGCAAAA	662
QY		866	TCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAAAATGAAAAACAAGAAATGTGTA	925
Db		663	TCTCAGCGATTTCAATCAAGTAAGGAGACTCCCATGAAAAATGAAAAACAAGAAATGTGTA	722
QY		926	TACCTACGGNACCAACCAAGGCGATAAGTGGCTTTAGNGATACCGATCTTTGGGCAGA	985
Db		723	CACCTACGGACCCACCAAGGCGATAAGTGGCTTTAGGAGATACCGATCTTTGGGCAGA	782
QY		986	AGTAGAACATGACTATACCACTATGCGGAAGAACTTAAATTTGGCGCGGTAAAACTAT	1045
Db		783	AGTAGAACATGACTATACCACTTATGGCGAAGAGCTCAATTTTGGCGCGGTAAAACTAT	842
QY		1046	CCGTGAGGGTATGGGTACAGACAAATGCCCTGTATGAAAAACACCTAGATTTAGTCATCAC	1105
Db		843	CCGTGAGGGTATGGGTACAGACAAATAGCCAGATGAAAAACACCTTAGATTTAGTCATCAC	902
QY		1106	TAAACGGATGATTTATCGACTACACCGGGATTTTCAAAAGCGCATTTGGGATTTAAACCG	1165
Db		903	CAACCGGATGATTTATCGACTACACCGGGATTTTAAAGCGCATTTGGTATTTAAAAATGG	962
QY		1166	CAAAATTCATGGCATTTGGCAAGCAGGAAACAAGGACATGCAAGATGGCGTAAAGCCCTCA	1225
Db		963	CAAAATTCATGGTATTTGGCAAGCGGGAACAAGACATGCAAGATGGCGTAAAGCCCTCA	1022
QY		1226	TATGGTCTGGGTGTGGGCA CAGAAGCACTAGCAGGGGAAGGTATGATTTATCCCGCTGG	1285
Db		1023	TATGGTCTGGGTGTGGGCA CAGAAGCACTAGCAGGGGAAGGTATGATTTATCCCGCTGG	1082
QY		1286	GGGAATCGATTCACACACCACTTCTCTTCTCCAACAANTTCCTCTACCGCTCTAGCCAA	1345
Db		1083	GGGGATCGATTCGCA CACCACTTCTCTCTCCCAACAANTTCCTCTACCGCTCTAGCCAA	1142
QY		1346	TGGGTGTACAAACCATGTTTGGAGCGGCA CAGGTCTGTAGATGGCACGAATGCGACTAC	1405
Db		1143	TGGTGTGTACAAACCATGTTTGGAGGTGGCA CAGGTCTGGTATGGCAAGNAATGCGACAC	1202
QY		1406	TATCACTCCGGGCAAAATGGAACCTTGCAACCGCATGTTTGGCGGCAAGAGATTTCTAT	1465

Db 1203 CATCACTCGGCGCAATGGAATCTGCACCGCATGTTGGCGCAGCTGAAGATTTCTAT 1262
 Qy 1466 GAATGTGGCTTTTGGGCAAGCAATAGCTCTAGCAAAAAACAATTTAGAACAAAGT 1525
 Db 1263 GAATGTAGCTTTTGGGCAAGCAATAGTTCTAGCAAAAAACAATTTAGAACAAAGT 1322
 Qy 1526 AGAAGCGGCGGATGTTTAAATTCATGAAGACTGGGGCACAACCAAGTGGCAT 1585
 Db 1323 AGAAGCGGCGGATGTTTAAATTCATGAAGACTGGGGCACAACCAAGTGGCAT 1382
 Qy 1586 CGATCACTGCTTGAGCGGCGAGATGAATAGATGCAAGTTTGTATCCACACCGATAC 1645
 Db 1383 CGATCACTGCTTGAGCGGCGAGATGAATAGATGCAAGTTTGTATCCACACCGATAC 1442
 Qy 1646 AGTCAATGAGCGAGTTATGTAGATGACACCTTAAATGAATGAACGGCGCGCCATCCA 1705
 Db 1443 GGTCAATGAGCGAGTTATGTAGATGACACCTTAAATGAATGAACGGCGCGCCATCCA 1502
 Qy 1706 TGCCTACACATTTAGGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGG 1765
 Db 1503 TGCCTACACATTTAGGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGG 1562
 Qy 1766 CGAGCTCAATATTTACCTCTCCACCCACCCCACTATTCCTATACCATTAATACGGT 1825
 Db 1563 CGAGCTCAATATTTACCTCTCCACCCACCCCACTATTCCTATACCATTAATACGGT 1622
 Qy 1826 TGCAAGACACTTAGACATGCTCATGACATGCCACCATCAGACAAAGCATCCCGAGGA 1885
 Db 1623 TGCAAGACACTTAGACATGCTCATGACATGCCACCATCAGACAAAGCATCCCGAGGA 1682
 Qy 1886 TTACAATTTCTCAAGCGGTATCCGCGCGCTCTATCCGGCTGAAGATGCTCCA 1945
 Db 1683 TTACAATTTCTCAAGCGGTATCCGCGCGCTCTATCCGGCTGAAGATGCTCCA 1742
 Qy 1946 TGATATGGTGTATCGGATGACAGCTCGGATTCGCAAGCATGGGGCTGAGGGA 2005
 Db 1743 TGATATGGTGTATCGGATGACAGCTCGGATTCGCAAGCATGGGGCTGAGGGA 1802
 Qy 2006 AGTGATTTCTCGAATTTGGGAGCTGGGATAGGATTAAGAAAGAAATTTGGTGAAGTTCC 2065
 Db 1803 AGTGATTTCTCGAATTTGGGAGCTGGGATAGGATTAAGAAAGAAATTTGGTGAAGTTCC 1862
 Qy 2066 TGAAGATGGCAAGATACGATATTTCCGATTAAGCGCTACATCTCCAAATACACTAT 2125
 Db 1863 TGAAGATGGGAGATACGATATTTCCGATTAAGCGCTACATCTCCAAATACACTAT 1922
 Qy 2126 CAACCCCGCTTTGACCAAGCGGTGAGGAGTATATCGGCTCTGTGGAAGGGCAAGAT 2185
 Db 1923 TAATCCCGCTTTGACCAAGCGGTGAGGAGTATATCGGCTCTGTGGAAGGGCAAGAT 1982
 Qy 2186 CGCGGACTTGGTGTGGATCTGCTTTTGGCGTAAACCCAAATCGTGATCAA 2245
 Db 1983 CGCGGACTTGGTGTGGATCTGCTTTTGGCGTAAACCCAAATCGTGATCAA 2042
 Qy 2246 AGCGGATGATGCTCTTCTGAAATGGGCGATTTAAAGCGCTCTGTGCGCACTCCCA 2305
 Db 2043 AGTGGAATGATGCTCTTCTGAAATGGGCGATTTAAAGCGCTCTGTGCGCACTCCCA 2102
 Qy 2306 ACCGGTTTATACCGCAAAATTTTGGGCAATTCAGCGCAAGCGGAAATTTGACACCGAT 2365
 Db 2103 GCGGTTTATACCGCAAAATTTTGGGCAATTCAGCGCAAGCGGAAATTTGACACCGAT 2162
 Qy 2366 CACTTTTGTTCGAAGTCG 2385
 Db 2163 CACTTTTGTCTCAAGCG 2182

RESULT 6
 ADQ37847
 ID ADQ37847 standard; DNA; 8407 BP.
 XX
 AC ADQ37847;

XX 07-OCT-2004 (first entry)
 XX H. bizozeronii urease gene cluster, ureABIEFGH.
 DE Urease; urease gene cluster; urease structural gene;
 XX Urease accessory gene; ureABIEFGH; Helicobacter bizozeronii infection;
 KW antibacterial; gene; ds.
 KW Helicobacter bizozeronii.
 OS US2004142343-A1.
 XX 22-JUL-2004.
 PN 12-AUG-2003; 2003US-00639273.
 PD 16-AUG-2002; 2002US-0404337P.
 PF (CHAN//) CHANG Y.
 XX (SIMP//) SIMPSON K W.
 PA (ZHUJ//) ZHU J.
 XX Chang Y, Simpson KW, Zhu J;
 PI WPI; 2004-533502/51.
 XX GENBANK; AF330621.
 DR Novel isolated nucleic acid molecule having urease gene cluster, and
 XX conferring on Helicobacter bizozeronii ability to produce urease, useful
 PT as vaccine for preventing disease in mammals infected by H.bizozeronii.
 XX Claim 2; SEQ ID NO 1; 40pp; English.
 CC The invention relates to an isolated nucleic acid molecule conferring on
 CC Helicobacter bizozeronii an ability to produce urease, where the nucleic
 CC acid molecule is a urease gene cluster comprising at least one urease
 CC structural gene and at least one urease accessory gene. The nucleic acid
 CC molecule is chosen from ureA, ureB, ureF, ureG, ureH and ureI. The
 CC invention also relates to an isolated protein encoded by the nucleic
 CC acid, a vaccine for preventing onset of disease in mammals infected by H.
 CC bizozeronii comprising a nucleic acid and a carrier, and an isolated
 CC antibody or its binding portion raised against the nucleic acid. The
 CC nucleic acids, proteins and antibodies are useful for vaccinating mammals
 CC against onset of disease caused by infection of H. bizozeronii, which
 CC involves administering the sequences. The sequences are useful for
 CC detecting H. bizozeronii in a sample of tissue or body fluids which
 CC involves providing a nucleic acid as an antigen, providing an antibody,
 CC or providing a nucleotide sequence as a probe in a nucleic acid
 CC hybridisation assay, contacting the sample with the antigen or the probe,
 CC and detecting any reaction which indicates that H. bizozeronii is
 CC present in the sample. This sequence represents the H. bizozeronii
 CC urease gene cluster, ureABIEFGH.
 XX
 SQ Sequence 8407 BP; 2264 A; 1937 C; 1969 G; 2237 T; 0 U; 0 Other;
 Query Match 39.4%; Score 1134.4; DB 12; Length 8407;
 Best Local Similarity 68.5%; Pred. No. 3.9e-227;
 Matches 1681; Conservative 0; Mismatches 741; Indels 32; Gaps 2;
 Qy 183 AAATTTAAACAAGAGGATTAAGTGAACCTCACACCAAGAGAGGCTTAAGCTCAACCAAC 242
 Db 2319 AAATTTGGTAGAGGATTTAGGATGAATTAACCCCTTAAGAGCTGGACAGCTCATGT 2378
 Qy 243 TATATTATGGCGGCAAGTGGCTAGAAAGCGAAAGAGAGGCTTAAGCTCAACCAAC 302
 Db 2379 TGCATTATGGCGGCAATTTGGCTAAAAACGCAAGCAAAATGCGCTTAAGCTAAATATA 2438
 Qy 303 CCGAAGCCATTGCTTACATTTAGTCCCATATTATGAGCGAAGCGCGCTGGAAGAAAAA 362
 Db 2439 CTGAGGAGTAGAGCCCTCATCATGTCGCCCATGTATGGAAGAGAGCCCGTCAGGTAAGAAAA 2498
 Qy 363 CCGTTGCCAGCTTATGGAAGAGTGCATGCACATTTTGAAGAAAAAGATGAAGTAATGCCCCG 422

Db 2499 GTGTGGCGATTTGATGCAAGAGCGAGGACACTTCTTAAAGCTGATGATGTCAATGCCCG 2558
Qy 423 GGGTGGGTAAATPATGGTTCCCGATCTAGGTGTAGAGCCACCTTTCTTGATGGTACGAAC 482
Db 2559 GTGTAGCCCATATGATCCACGAAGTGGGATTAAGACTAACTTCCCTGATGGACAAAC 2618
Qy 483 TTGTAACTGTGAATTTGGCCCATCGAAACAGATGAGCACTTCAAGCGGGCGAAGTGAAT 542
Db 2619 TGGTAAACCATTCATACCCCGTTGAAGATGTGGGCAATAATGGCTCCGGTGAAGTGA 2678
Qy 543 TTGTGTGGATAAAGACATCGAGCTCAATGCAGGCAAGAGAGTAAACCGAACTTGAGGTTA 602
Db 2679 TTTTGAAGAAACGAAGACATCACTTTGAATGCAGGCAAGCAACGACCACTTTAGAGTGC 2738
Qy 603 CTAATGAAGGCCCTAAATCTTGTGATGTGGGTAGCCATTTCCACTTCTTTTGAAGCTAACA 662
Db 2739 ATAACAAAGCGATCGCCCGTGCAGTGGCTCCCACTTCCACTTCTTTGAAGTGAATA 2798
Qy 663 AGGCACCTAAATTCGATCGTGAAGAAAGCCTATGGCAAAACGCTAGATATCCCTCTGCGCA 722
Db 2799 AGCTTTTGGAAATTTGATCGTGAAGAAAGCCTATGGCAAAACGCTAGACATTTGCTTCTGAA 2858
Qy 723 ACAGCTACGATTTGGGCGAGGACAAACCCGAAAGTGCAGTGTGATTTCTTGTGGCA 782
Db 2859 CCGCTGTGCGCTTTGAACCCGCTGAGAAAGAAACCGTGAATTTGATTCAAATTTGGCGGTA 2918
Qy 783 GTAAAAAAGTGAATGGCATGAACGGCTTTGTGAATAACATCGCGGATGAACGCCATAAAC 842
Db 2919 ACCAAGCATTTACGGCTTTAACTCTCTTGTGGATCGCAAGCGGATGATGATGCAAAA 2978
Qy 843 ATAAAGCGCTTGAACAGCGGAAATCTCACGATTT----- 877
Db 2979 AACTTGTCTCTCAACGGCGCAAGAAACATGCTTTGTGTTGTGAATTTGGCGTTCGCGATA 3038
Qy 878 ----ATCAAGTAAAGAGACTCCCATG----AAATGAAAAAACAAGATATGTAATACCT 930
Db 3039 AAAAAATAGGAAGGACAAATCCGATGAAGAAATCTCTCGAAAGAAATGTTTCTATGT 3098
Qy 931 ACAGACCCACCAAGGCGATAAAGTGCCTTTAGAGATACCAGTCTTTGGGCGAGAAGTAG 990
Db 3099 ATGAGACCACTACCGGCGATAAAGTGAATTTGGGCGATACCAGCTGATCTTAGAGTGC 3158
Qy 991 AACATGACTATACCACTATGCGGAAGAACTTAAATTTGGGCGGTTAAAACTATCCGTG 1050
Db 3159 AACATGACTGACCACTTATGCGGAAGAAATTAAGTTTGGTGGCGTAAAAACCATTTCCG 3218
Qy 1051 AGGATATGGGTACAGACATAGCCCTGATGAACACCCCTAGATTTAGTCACTAAGC 1110
Db 3219 ATGGGATGGCAAAAACCAACAGCCCCAGCAGCCACGAACTCGATCTTGTGCTCACTAAGC 3278
Qy 1111 CGATGATTTATCGACTACACCGGGATTTTACAAGCCGACATTTGGGATTAAGAAACGGCAAAA 1170
Db 3279 CCCTGATCGTGGATTAACCGCGATTTTATAAGCGGATTTTGGCATTAAGAAATGGCAAAA 3338
Qy 1171 TCCATGGCATTTGGCAAGCGCAATAAGACATGACAGATGGCGTTTGGCAACATCTTT 3398
Db 1231 TCCTGCGGTGGGCAACAGACATAGCAGGGGAAGTATGATTAATACCGTGGGGAA 1290
Qy 3399 GCGTGGGCGCTTACTGAGGCTTTGGCGCTGAAGGGCTGATTTGTTACAGCTGTGGGA 3458
Db 1291 TCGATTCACACCCACTTCTTCTCCACAACTTCCCTACCGCTACCGCTAGCCATGGCG 1350
Qy 3459 TTGACACCACTCACTTATTTCTCCCAACAAATCCCAACAGCATTTGCCAGCGGA 3518
Db 1351 TTACAAACCATGTTGGAGCGGCAAGCTCTGTAGATGGCAACGAATCGGACTACTATCA 1410
Qy 3519 TCACAAACATGATTTGGTGGGGGAACAGGTCCAGCTGATGGGACTTAACCGGACTACATCA 3578
Db 1411 CTCGGGCAAAATGGAACCTTGCACCGCATTTGTGCGGCGACGAGAGATTTCTATGAATG 1470

Db 3579 CTCGGGGCGCTGGAACTTAAAAACATGCTCGTGGCTCTCGAAGAAATATGCCATGAAC 3638
Qy 1471 TGGGCTTTTGGGCAAGCAATAGCTCTAGCAAAAAACAACTTTGAGAAACAGTAGAAG 1530
Db 3639 TGGGCTATTTGGGTAAAGGAATGTGTCTTATGAACCCCTCCCTGGTCGATCAACTCGAAG 3698
Qy 1531 CGGGCGCATTTGGTAAAAATTTGAATGACATGCGGGCAACACCAAGTGCATCGATC 1590
Db 3699 CTGGAGCCATTTGGCTTTAAAAATCCAGAAAGATGCGGGTAGCACCTGCGAGCATATACC 3758
Qy 1591 ACTGCTTCAGCGCTGCGAGATGAATACGATGTGCAAGTTTGTATCCACACGATACAGTCA 1650
Db 3759 ATTGCTTGAATGTGCTGACAAATACGATGTGCAAGTGTCTATCCACACGATACCTTGA 3818
Qy 1651 ATGAGCGAGTTATGTAGATGACACCTAAATTCGAATGAACGGCGCGCATCCATGCTCT 1710
Db 3819 ATGAAGCGGCTGTGTGGAAGACACTTTTCAAGCCATTTCTGGGCGCACTATCCACACTT 3878
Qy 1711 ACCAATTTGAGGGAGCGGTGAGGACACTCACCTGATGTTTATCACCATGGCAGGCGAGC 1770
Db 3879 TCCACACTGAAGGTGCTGTGGTGGCGGCGACGCTCCGATGTCTAAGATGTCTGTGGCGAAT 3938
Qy 1771 TCAATATTTCTACCTCTCCACACCCCACTATTTCCCTATACCATTAATACGTTGCGAG 1830
Db 3939 TTACATCTCTCCAGCTTCTACCAACCCCACTTCTTCCACCTGGAATACAGAAAGCCG 3998
Qy 1831 AACCTTAGACATGCTCATGACATGCCACCATAGACAAACGCGCATCCGCGAGGATTTAC 1890
Db 3999 AACACATGACATGTTGATGGTGTGCCACCATTTGATATAAAACATCAAGAAAGATGTCC 4058
Qy 1891 AATTTTCTAAAGCCGATTCGCCCGCGGCTCTATGCGGCTGAAGATGTCTCCATGATA 1950
Db 4059 AGTTTGTGATTTCTAGGATTCGCCCGCAACCATGCGCGCTGAGGCAAACTCCACGATA 4118
Qy 1951 TGGGTGATCGCGATGACAACTCGGATTTGCAAGCAATGGCGCTGCGAGGCGAAGTGA 2010
Db 4119 TGGGATTTTCTCTATACCACTCTGACTCCCAAGCGATGGCGCTGATGAGGAGGTCA 4178
Qy 2011 TTCTCTGAACTTTGGCAGACTGCGGATAGAAATAAAAAAGAAATTTGGTAAAGTCTCTGAAAG 2070
Db 4179 TCACCGGCACTTTGGCAACAGCGGACAAAAACAAAAAGAAATTTGGTTCGCTTGCGTGGG 4238
Qy 2071 ATGGCAAGATAACGATTAATTTTCGCAATTAAGCGGTACATCTCCAAATACATATCAACC 2130
Db 4239 AAAAAGGCGATAATGACAACTTCCGCACTCAAGCGCTTACATTTCCAAATACCATCAACC 4298
Qy 2131 CGCTTTGACCCACCGCGTGGGAGTATATCGGCTCTGTGGAAGAGGCGAGATCGCG 2190
Db 4299 CGCTTAITGGCACACGCGATTTCTGAATATGTGCGCTCTGTAGAAGTGGGCAAAATTCGCG 4358
Qy 2191 ACTTGGTGTGGATCTGCTCTTTTGGGCGTAAAAACCCAAATCTGTGATCAAAAGCG 2250
Db 4359 ATTTGGTCTTTGGAGTCTGCTGTTTGGCATTAACCCAAACATGATCATCAAGGCG 4418
Qy 2251 GTATGGTGTCTTCTGAAATGGCGATTTTAAACGCTGTGTGCCACTTCCCCAAACCGG 2310
Db 4419 GATTCATCGCACTTCTCAATGGGCGATGCCAATGCGCTCTATCCCACTTCCCCAAACCGG 4478
Qy 2311 TTTATTCGCGAAATGTTTGGGCGATCAGCGCAAGCGGAAATTTGACACAGCATCACTT 2370
Db 4479 TGTATTCGCGAAATGTTTGGGCGATGATGCGGCAATGTTTGAACCAATATCACTT 4538
Qy 2371 TTGTTTCAAAAGTCTGCTATGAAATTTGGGCGTGAAGAAAGCTGGGCTTAGAGCGCAAG 2430
Db 4539 TTGTATCCAGTGGCTTATGACACCGCATTTAAGAAAGTGTGGCTTGCAGAAAGTGG 4598
Qy 2431 TTCTACCGGTCAAAACTGCCGTAAACATCAACCAAGAAAGACTTCAAGTTTCAACGACAAA 2490
Db 4599 TTTTTCGAGTTTAAAAAATGCCCGCAACATCAACCAAAAAAGACCTCAAAATTCACGATGTTA 4658
Qy 2491 CGGCAAAATACCGTGCATCCGAAACCTTCCAGGTCTTTGTAGATGGCAAACTCTGCA 2550
Db 4659 CGGCACATCGAAGTCAATCTCTGAAACCTTACAAAGTTTAAAGTGGATGGCAAGAGGTTA 4718

QY 987 GTAGACATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGTAAACTATC 1046
Db 915 GTAGACATGACTACACCATTTATGGCGAAGAGCTTAAATTTGGCGGGTAAACCCCTA 974
QY 1047 CGTGAAGGTATGGGTACAGCAATAGCCCTGATGAAACACCCCTAGATTTAGTATCACT 1106
Db 975 AGAAGGCATGAGCCCAATCTAAACCCCTAGCAAGAGAGTTGGATTTAATATCACT 1034
QY 1107 AACCGCATGATATGACTACACCGGGAATTACAAAGCCGACATTTGGGATTTAAACCGC 1166
Db 1035 AACCGTTTAAATCGTGGATTTACACCGGATTTATAAAGCGGATTTGGTATTAAAGATGGC 1094
QY 1167 AAAATCATGCAATTTGGCAAGCAGCAAAACAGGACATGCAAGTGGCGTTAAGCCCTCAT 1226
Db 1095 AAAATTCGCTGGCAATTTGGTAAAGCGGTTAAACAGACATGCAAGTGGCGTTAAACCAAT 1154
QY 1227 ATGTCGTCGGGTGGGACAGAACCACTAGCAGGGGAAGGTATGATTTATACCGCTGG 1286
Db 1155 CTTAGCGTAGTCTGCTACTGAGCCCTTAGCCGTTGAAGTTTGATCGTAACGGCTGGT 1214
QY 1287 GGAATCGATTACACACCCACTTCCTTTCTCCACAACTTCCCTACCGCTCTAGCCAAAT 1346
Db 1215 GGTATTGACACACATCCACTTCATTTTCAACCCCAACAAATCCCTACAGCTTTTGCAGC 1274
QY 1347 GCGGTTACACCATGTTGGAGGGGCGACAGTCTCTGTAGATGGACGAATGCGCACTACT 1406
Db 1275 GGTGTAAACCAATGATGTTGGTGGAAACCGGCTCTGCTGTGGCACTTAATGCGCACTACT 1334
QY 1407 ATCACTCGGCGCAATGAACTTGACACCGCATGTTGCGGCGACAGAGAGTATTCATATG 1466
Db 1335 ATCACTCAGGCGAGAGAAATTTAAATGATGCTCAGAGCGGCTGAAGAAATATTCATATG 1394
QY 1467 AATGTCGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACACTTTGTAGAAACAAGTA 1526
Db 1395 AATTAGTCTTTGGCTAAAGGTAAACGCTTCTACGATGCGAGCTTAGCCGATCAATTT 1454
QY 1527 GAAGCGGCGGATGTTGTTTAAATGCAATGAAGACTGGGGCAACACCAAGTGGCATC 1586
Db 1455 GAAGCGGTCGATGTTGGCTTTAAATTTACGAAGACTGGGGCACCACTCTCTCTGCAATC 1514
QY 1587 GATCACTGCTTGAGGCTGGCAGATGAATACGATGTCGAAGTTTGTATCCACCGATACA 1646
Db 1515 AATCATGCGTTAGATGTTGGCGACAAATACGATGTCGAAGTGGCTATCCACACAGACT 1574
QY 1647 GTCAATGAGGCGAGTTATGTAGATGACACCCCTTAAATGCAATGAACGCGCGCCATCCAT 1706
Db 1575 TTGAATGAAGCGGTTGTGTAGAGACACTATGCTGCTATTTGCTGGAGCACTATGAC 1634
QY 1707 GCCTACCACTGAGGAGCGGTTGGAGACACTCACTGTATGTTATCACCATGGCAGGC 1766
Db 1635 ACTTTCCACACTGAAGCGCTGGCGGCGACACGCTCTGTATTTTAAAGTAGCCGGT 1694
QY 1767 GAGTCATATTTCACTCTCTCCACCAACCCCACTATTCCTATACCATTAATACGTT 1826
Db 1695 GAACACAACTTTCTCCGCTTTCCACTAACCCCACTCTTCCCGTGAATACAGAA 1754
QY 1827 GCAGAACACTTAGACATGCTATGATGACATGCAACCACTAGACAAACGATCCGCGAGAT 1886
Db 1755 GCAGAGCACTGACATGCTTATGTTGTGCGCACTTGGATTAAGCAATTAAGNAGAT 1814
QY 1887 TTACAAATTTTCTCAAAGCGTATCCGCGCCGCTCTATCGCGCTGAAGATGTGCTCCAT 1946
Db 1815 GTTCAGTTTCGCTGATTTCAAGGATCCGCTCAACCAATTTGGCGCTGAAGACACTTTGCAT 1874
QY 1947 GATATGGTGTATCGGATGACAGCTCGGATTCGCAAGCAATGGCGGCTGCGAGCGAA 2006
Db 1875 GACATGGGGATTTTCTCAATCACCACTGCTGACTCTCAAGCGATGGCGGCTGTGGGTGAA 1934
QY 2007 GTGATTCCTCGAATTTGCGAGCTCGGATAGATAAAGAAATTTGTAAGCTTCTCT 2066
Db 1935 GTTATCACTAGAACTTTGGCAACAGCTGACAAAAAAGAAATTTGGCGGCTTTGAAA 1994
QY 2067 GAAGATGGCAAGATTAACGATAAATTTCCGCAATTAAGCGCTATCATCTCCAATAACACTATC 2126

Db 1995 GAAGAAAAAGGCGATAACGAACTTTCAGGATCAAAACGCTACTTGTCTAAATACACCAT 2054
QY 2127 AACCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGTGAAGAGGCGCAAGATC 2186
Db 2055 AACCCAGCGATCGCTCATGGGATTAGCGAGTATGTAGTTTCAGTAGAAGTGGGCAAGTG 2114
QY 2187 GCGGATTTGGTGTGTGAAATCTCGCTTTTGGGCTTAAACCCCAAAATCGTGATCAAA 2246
Db 2115 GCTGACTTTGGTATTGTGGAGTCCAGCATTTCTTTGGCGTGAACCCCAACATGATCATCAA 2174
QY 2247 GCGGATTTGGTGTCTCTCTGAAATGGCGGATTTCAACGCGCTGTGTGCGCACTCCCCAA 2306
Db 2175 GCGGATTTGTCGTTAAGCCAAATGGCGGATGGAACGCTTCTATCCCTACCCCAAA 2234
QY 2307 CCGGTTTATTACCGGAAATGTTGGGCAATCACGGCAAGCGGAAATTTGACACCAAGATC 2366
Db 2235 CCGGTTTATTACAGAGAAATGTTGCTCATCATGTTAAGCTTAATACGATGCAACATC 2294
QY 2367 ACTTTGTTTCAAAAGTGCCTTATGAAATGCGGTGAAAGAAAGCTGGGCTTTAGAGCGC 2426
Db 2295 ACTTTGTTGCTCAAGCGGCTTATGACAAAGCAATTAAGAAAGAAATTAGGACTTGAAGA 2354
QY 2427 CAAGTTCTACCGGTCAAAAACCTGCGTACATCACCAAGAAAGACTTCAAGTTCAAGGAC 2486
Db 2355 CAAGTTGTTGCGGTTAAAAATTTGCAGAAATATCATCTAAAAAAGACATGCAATTTCAACGAC 2414
QY 2487 AAAACGGCAAAATCACCGTCTGATCCGAAACCTTCGAGGCTTTTGTAGATGGCAAACTC 2546
Db 2415 ACTACTGCTCATATTGAAGTCAATCTGNAACTTACCATGTGTTCTGTGATGGCAAGAA 2474
QY 2547 TGCACTTTAAACCCACTCGCAAGTGCCTTAGCCAGCGGTACACTTTTCTTAGG 2604
Db 2475 GTAACTTCTAAACCGCAATAAAGTAGAGCTTTGGCGCACTCTTTAGCATTTTCTAGG 2532

RESULT 8
AAQ90180
ID AAQ90180 standard; DNA; 2619 BP.
XX AAQ90180;
XX 25-MAR-2003 (revised)
DT 02-NOV-1995 (first entry)
XX Helicobacter felis urease urea/ureB operon.
XX Urease; UreA; UreB; vaccine; Helicobacter infection; UreI;
KW heat shock protein; ss.
XX Helicobacter felis.
XX Key Location/Qualifiers
RBS 31..37
FT /*tag= b
FT /note= "ureA Shine-Dalgarno site"
CDS 43..756
FT /*tag= a
FT /EC_number= "3.5.1.5"
FT /note= "UreA"
RBS 756..759
FT /*tag= d
FT /note= "ureB Shine-Dalgarno site"
CDS 766..2475
FT /*tag= c
FT /EC_number= "3.5.1.5"
FT /note= "UreB"

XX W09514093-A1.
XX 26-MAY-1995.
XX 19-NOV-1993; 93WO-EP003259.

XX 19-NOV-1993; 93WO-EP003259.
 XX (INSP) INST PASTEUR.
 XX (INRM) INST NAT SANTE & RECH MEDICALE.
 XX Labigne A, Suerbaum S, Ferrero R;
 XX WPI; 1995-200383/26.
 DR P-PSDB; AAR74336, AAR74337.
 XX
 XX Immunogenic composition against Helicobacter infection - also gene
 PT fragment(s) and protein(s) from Helicobacter urease gene cluster and heat
 PT shock protein(s).
 XX
 PS Claim 10; Fig 3; 128pp; English.
 XX
 CC The sequence encodes urease UreA and UreB proteins, which are components
 CC of a novel immunogenic composition capable of inducing protective
 CC antibodies against Helicobacter infection. The composition may include
 CC the UreA or UreB proteins, a urease- associated heat shock protein
 CC (AAR74338-39) or the UreI protein (AAR74340). The composition is used to
 CC prepare a vaccine for humans or animals, especially against H. pylori and
 CC H. felis. Antibodies against the proteins may be used for treating
 CC Helicobacter infection, and primers/probes to the DNA sequence may be
 CC used for detection of Helicobacter infection. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 XX Sequence 2619 BP; 772 A; 569 C; 665 G; 613 T; 0 U; 0 Other;
 Query Match 35.0%; Score 1008.8; DB 2; Length 2619;
 Best Local Similarity 66.2%; Pred. No. 5.5e-201;
 Matches 1634; Conservative 0; Mismatches 792; Indels 42; Gaps 8;
 QY 175 TGCCTTTAAATTAACACAGAGGAGTAATAGTGAAGTCAACCCAGAGCAAGAAA 234
 DB 13 TACCAATAGAAATTCATAGAGGAGTTAGATGAATCAACGCTTAAGAACTAGACAA 71
 QY 235 GTTCTGTGATATATATGCGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCT 294
 DB 72 GTTANTGCTCCATATATGCGGCGAGATTTGGCAGAGAACGCTTGGCGGTGTGAACT 131
 QY 295 CAACCAACCCGAGCCATTTGCTTACATAGTGGCCATATATAGCAGAGCGCCGCTGG 354
 DB 132 CAATTACACGAGCGGTGCGGCTCATTAGCGGCGTGTGATGCAAAAGGCGCGTATGG 191
 QY 355 AAAAAAACCGTGGCCAGCTTATGGAAGAGTGATGCACTTTTGAAGAAAGATGAGT 414
 DB 192 TAAATAAAGCGTGGCGGATTTGATGCAAGAGGCGAGGACTTGGCTTAAAAAAGAAAATGT 251
 QY 415 AATCCCGGGTGGGTAAATATGTTCCGATCTAGGTGTAGAACCCACTTTTCCTGTATGG 474
 DB 252 GATGACGGGTAGCAAGCATGATTCATGAGTGGGATTTGAAGTAACTTCCCGCATGG 311
 QY 475 TACGAAACTTGTAACTGTGAATTTGGCCCATCGAACCCAGATGAGCACTTCAAGCGGCGA 534
 DB 312 AACCAAGCTTGTAACTATCCACATCCCGGTAGAGGATAATGGCAATTTAGCCCCGGGA 371
 QY 535 AGTGAATTTGGTGGGATTAAGACATCGAGCTCAATGCGAGGCAAGAGTAACCGAACT 594
 DB 372 GGT---CTTCTTAAAAATGAGGACATTAATTAACGCGCGGCAAGAGCAATTAGCTT 428
 QY 595 TGAGGTACTAATCAAGGCGCTAAATCTTGTGATGCGGTAGCCATTTCCACTTCTTGA 654
 DB 429 GAAAGTGAATAAAGCGCATGCTGTCGAGGTGGGATCAATTTCCACTTCTTCGA 488
 QY 655 AGCTTAAACAGGCACTAAATTCGATCGTGAAGAACGCTATGCGCAAGCGCTAGATATTC 714
 DB 489 AGTGAATAAGCTCTTGACTTGTGATTCGCGCAAAAGCTTTTGCAACGCTAGACATTGC 548
 QY 715 CTCTGGCAACGCTAGCAATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTCT 774
 DB 549 ATCTGGAAACAGCGTGGCTTTGAAACCCCGGGAGGAAAAAAGTGTGGAATCTATTGACAT 608

QY 775 TGGTGGCAGTAAAAAGTGAATGGCATGAACGGGCTTGTGAATAACATCGCGAT---GA 831
 DB 609 CGGCGGGAATAAGCGCATCTATGGCTTTAAATTTCTTTGGTGGATCGCAAGCCGATGCCGA 668
 QY 832 AGCCCATAAAC-----ATAAAGCGCTTGAACA-----GGCGAAATCTCAACGG 873
 DB 669 TGGTAAAAAACTCGGCTTAAACGCGCTAAGAAAAAGGTTTGGGTCTGTAACTCGGG 728
 QY 874 ATTT--ATCAAGTAAGGAGACTC-----CCATGAAATGA---AAAAACAAGA 916
 DB 729 TTGTGAAGCGACTAAAGATAAACAATAAGGAAAAACCATGAAAAAGATTTCACGAAAGA 788
 QY 917 ATATGTAATACTACCTACGACCCACCAAGCGGATAAAGTGGCTTAGGAGATACCGATCT 976
 DB 789 ATATGTTTCTATGTATGTGTCCTCACTACCGGGGATCGTGTAGACTCGCGGACACTGATTT 848
 QY 977 TTGGGCGAGAAGTAGAACATGACTATACCACTATGCGGAAGCACTTAAATTTTGGCGGG 1036
 DB 849 GATCTTGAAGTGGAGCATGATTTGCACCACTTATGTGTGAAGAGATCAAAATTTGGGGCGG 908
 QY 1037 TAAAACTATCCGTGAGGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTT 1096
 DB 909 TAAACTATCCGTGATGGATGAGTCAACCAATAGCCCTAGCTCTTATGAATTAGATTT 968
 QY 1097 AGTCACTACTAAACGCGATGATTTATCGACTACACCGGGATTTACAAAGCGGACATTTGGGAT 1156
 DB 969 GGTGCTCACTAAACGCGCTCATTTGTGACTATATCGGGCAATTTACAAAGCGGACATTTGGGAT 1028
 QY 1157 TAAAAACGCAAAATCCATGGCATTTGGCAAGCGAGCAAGCAATGCAAGTGGCGT 1216
 DB 1029 TAAAGACGCGAAGATTGACAGGCAATTTGGCAAGCGAGCAATAGGCAATGCAAGATGGCGT 1088
 QY 1217 AAGCCTCATATGCTGCTGGGTGTGGGACAGACACTAGCAGGCGGAAGGTATGATTAT 1276
 DB 1089 AGATAATATCTTCCGTAGGTCTGCTACAGAGGCTTTGGCAGCTGAGGGCTTGATTGT 1148
 QY 1277 TACCGCTGGGGAAATCGATTACACACCCACTTTCCTTTCTCCAACAATTCCTCTACCCG 1336
 DB 1149 AACCGCTGTGGCATCGATAGCATATTTCACTTTATCTCTCCCAACAATCCTACTGC 1208
 QY 1337 TCTACCAATGGCGTTTACCAACATGTTGGAGCGGCGACAGGTCTCTGTAGATGGCAGAA 1396
 DB 1209 TTTTCCAGCGGGGTATCAACCATGATTGGAGGAGGACAGGACCTGCGGATGGCAGAA 1268
 QY 1397 TCGCACTACTATCACTCCGGGCAATGGAATTTGCACCCATGTTGCGGCGGACAGCAAGA 1456
 DB 1269 TCGGACCACTCACTCCCGGACGCGTAACTCTAAAAAGTATGTTGCGGTGACGCGCAAGA 1328
 QY 1457 GTATTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTGT 1516
 DB 1329 ATACCCCATGATCTAGGCTTTTGGCTAAGGGGAATGTGCTTACGAACCTCTTTACG 1388
 QY 1517 AGAACAACTGAGAGCGGCGGATTTGGTTTAAATTTGATGAAGACTTGGGGCAACACCC 1576
 DB 1389 CGATCAGATTGAAGCAGCGGCGGATTTGTTTAAATTCACGAAAGACTTGGGGAAGCAGCAC 1448
 QY 1577 AAGTGGATCGATCACTGCTTGGAGCTGGCAGATGATACGATGTGCAAGTTTGTATCCA 1636
 DB 1449 TGCAGCTATTACCACTGCTCAATGTGCGCGATGAATACGATGTGCAAGTGGCTATCCA 1508
 QY 1637 CACCGATACAGTCAATGAGGCAAGGTATGTAGATGACACCTTAAATGCAATGAACGGGCG 1696
 DB 1509 CACCGATACCTTAAAGAGGCGGCTGTGTAGAGACACCTTAGAGCGGATTTCCCGGGCG 1568
 QY 1697 CGCCATCCATCCCTACCACTTGAAGGAGCGGGTGGAGACACTACCTGATGTTATCAC 1756
 DB 1569 CACCATCATACCTTCCACACTGAAGGGGCTGGGGGTGGACACGCTCCAGATGTTATCAA 1628
 QY 1757 CATGGCAGCGAGCTCAATATTTCTACCTCTTCCACCCCGGCACTATTTCCCTATACCAT 1816
 DB 1629 AATGGCAGGGAAATTTAAACATTTCTACCGGCTCTACTAACCGGACCAATTTCTTTCCAA 1688

QY 475 TACGAAACTTGTAACTGTGAATTTGGCCATCGAACCCAGATGAGCACTTCAAAGCGGGCGA 534
DB 312 AACCAAGCTTGTAACTATCACACTCGGTAGAGATAATGCGCAATTTAGCCCGCGGCGA 371
QY 535 AGTGAATTTGGTTGCGATAAAGACATCGAGCTCAATTCAGGCGAAAGAAAGTAACCGAACT 594
DB 372 GGT---CTTCTTAAATAATGAGGACATTAATTAATTAACGCGCGCAAAAGAACCAATTAGCTT 428
QY 595 TGAGGTTACTTAATCAAGGGCTTAATCTTGCATGTGGGTAGGACATTTCCACTTCTTTGA 654
DB 429 GAAAGTGAATAAAGCGCATCGTCTGTGTCAGGTGGGATACATTTTCCACTTCTTTCGA 488
QY 655 AGCTTAACAGGACCTAAATTTGCATGTGTAAGAAAGCCTATGCGAAAGCCCTAGATATTTCC 714
DB 489 AGTGAATAGCTTTGACATTTGCATTCGCGCAAAAGCTTTTGCAAAAGCCCTAGACATTCG 548
QY 715 CTCTGCAACACATCGCATTTGGGGGAGGACAAACCCGCAAGGTGCAAGTTGATTTCTCTCT 774
DB 549 ATCTGGAACAGCGGTGCGCTTTGAAACCGGGAGGAAAGGTGTGGAACATCAITGACAT 608
QY 775 TGGTGGCAGTAAATAAGTATTTGCATGAACCGGCTTTGTGAATTAACATTCGCGGAT---GA 831
DB 609 CGCGCGAATAAAGCGCATCTATGCTTTAAATCTTTGTTGGTATCGCCAAAGCCGATGCCGA 668
QY 832 ACGCCATAAAC-----ATAAGCGCTTGACAA-----GGCGAATCTCACGG 873
DB 669 TGGTAAATAAAGCTCGCTTTAAACCGCGCTTAAAGAAAGGTTTGGGCTGTGTAAGTCCGG 728
QY 874 ATTT---ATCAAGTAAGGAGACTC-----CCATGAAATATGA---AAAAACAAGA 916
DB 729 TTGTGAAGCGACTTAAGATTAACAATGAAGAAACCATGAAAGATTTTACGAAAGA 788
QY 917 ATATGTAAATACCTTACGGAACCCACCAAGGCGATAAAGTGGCTTTAGGAGATACCGATCT 976
DB 789 ATATGTTTCTATGATGTCTCCACTACCGGGGATCGTGTAGACTCGCGGACACTGATTT 848
QY 977 TTGGGCGAAGTAGAATGATGATACCACTATGCGGAAGAACTTAAATTTGGCGGGGG 1036
DB 849 GATCTTTAGAAAGTGAGCATGATTTGACCACTTATGTTGTAAGAGATCAAAATTTGGGGCGG 908
QY 1037 TAAACCTATCCGTGAGGATGAGTGTGAGCAAGCAATAGCCCTGATGAACCAACCTTAGATTT 1096
DB 909 TAAACCTATCCGTGAGGATGAGTGTGAGCAAGCAATAGCCCTGATGAAATAGATTT 968
QY 1097 AGTCATCACTAAACCGGATGATTTACACTACACCGGATTTTAAAGCCGACATTTGGGAT 1156
DB 969 GGTGCTCACTAAACCGCTCACTGTGACTATACCGGATTTTAAAGCCGACATTTGGGAT 1028
QY 1157 TAAAAACGGCAAAATCGATGGCAATTTGGCAAGGAGGAAACAAAGACATGCAAGATGCGGT 1216
DB 1029 TAAAGACGGCAAGATTTGACAGGCATTTGCAAGGCGAGCAATAAGGACATGCAAGATGCGGT 1088
QY 1217 AAGCCCTCATATGTTGTTGGTGTGGGCAAGACACTAGCAGGGAAGGTATGATTTAT 1276
DB 1089 AGATAATAATCTTTGGGTAGGCTCTGCTACAGAGGCTTTTGGCAGCTGAGGCTTTGATTTG 1148
QY 1277 TACCGTGTGGGGAATCGATTTACACACCACTTCTTCTTCCCAACAAATTTCCCTACCGC 1336
DB 1149 AACCGTGTGGGATCGATACGCAATTTCACTTTATCTCTCCCAACAAATTTCCCTACTGC 1208
QY 1337 TCTAGCAATTTGGGTTTACACCAATTTTGGAGCGGCGACAGGCTCTGTAGATGSCAGAA 1396
DB 1209 TTTTGGCAGCGGGTTTACACCAATGATTTGGAGGAGGACAGGACCTGCGGATGGCAGAA 1268
QY 1397 TGGCACTATATCACTCGGCGGCAAAATGGAACCTTGGCAACCGCATGTTGGCGGAGCAGAA 1456
DB 1269 TGGCAACCACTCACTCCCGGACCGCTATCTTAAAGAGATGTTGTTGGTGCAGCGGAGAA 1328
QY 1457 GTATTTCTATCAATGCGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAAACCACTTGT 1516
DB 1329 ATACGCCATGAATCTAGGCTTTTGGCTTAAGGGGAATGTGCTTACGAACCCCTCTTTACG 1388
QY 1517 AGAAAGTAGAAGCGGGCGGATTTGGTTTAAATTTGCAATGAAGACTGGGGGCAACACACC 1576

DB 1389 CGATCAGATTTGAAGCAGGGCGATTTGGTTTAAAAATCCACGAGACTGGGGAAGCACACC 1448
QY 1577 AAGTCGATCGATCACTGCTTGGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCA 1636
DB 1449 TGCAGCTATTACCACTGCTCCTCAATGTGCGCGATGAATACGATGTGCAAGTGTCTATCCA 1508
QY 1637 CACCGATACAGTCAATGAGGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGG 1696
DB 1509 CACCGATACCCCTTAAAGGAGGGGGCTGTGTAGAAGACACCCCTAGAGGCGATTTGCCGGGG 1568
QY 1697 CGCCATCATGCTACCACTTGAAGGAGCGGTGAGAGCACTACACCTGATGTTATCAC 1756
DB 1569 CACCATCATTACCTTCCACATGAAGGGGTGGGGTGGACAGCTTCCAGATGTTATCAA 1628
QY 1757 CATGCGAGCGAGCTCAATATTTCTACCTCTCTCCACCCACCCCTATTTCCCTATACCAT 1816
DB 1629 AATGCGAGGGGAATTTAAACATTTCTACCCGCTCTACTAACCCGACCATTTCTTCCACAA 1688
QY 1817 TAATACGTTTGCAGAACACTTAGACATGTCTATGACATGCCACCACTAGACAAACGCAT 1876
DB 1689 AAACACTGAAGCCGAGCACATGATTTATGTTGTGTCACCACTTGGATAAAAGTAT 1748
QY 1877 CGCGAGGATTTACAAATTTCTCAAGCGGTATCCGCCCGGCTCTATCGCGCTCAAGA 1936
DB 1749 CAAGGAAGATGTGCAAGTTTGGCGATTCGAGGATTCGCCCCCAAACTATCGCGCTGAAGA 1808
QY 1937 TGTGCTCATGATATGGGTGTGATCGCATGACAAAGCTCGGATTCGCAAGCAATGGGGCG 1996
DB 1809 CCAACTCATGACATGGGATCTTTCTATCACAGCTCCGACTCTCAGGCTATGGGACG 1868
QY 1997 TGCAGGGAAGTGTATCTCGAACTTTGCGAGCTGCGGATAGAATAAAAAAGAAATTTGG 2056
DB 1869 CGTAGCGAGGTGATCACACGCTTTGCGACAGACAGACAAAAACAAAAAGATTTGG 1928
QY 2057 TAAGTCTCTGAAGTGCAGAAAGATAACGATATTTTCGCATTAAGCGCTACATCTCAA 2116
DB 1929 GCGCTTGAAGAGGAAAAAGGCGATTAACGCAACTTTCGCACTCAACCGCTACTCTCAA 1988
QY 2117 ATACACTATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGA 2176
DB 1989 ATACACCATCAACCCCGGATCGCGATGGATTTCTGACTATGTTGGCTCTGTGGAAGT 2048
QY 2177 GGGCAAGATCCCGACTTTGGTGTGGAATCTCTGCTTTTGGCGTAAAAACCCAAAT 2236
DB 2049 GGGCAATATACCGGACCTCGTGTCTTTGGAGTCCGCTTTCTTTGGCATTAAGCCCAATAT 2108
QY 2237 CGTATCAAGCGGATGTTGGTGTCTCTGGAATGCGCGATTTCTAAGCGCTGTGTGCC 2296
DB 2109 GATTTAAGGGCGGATTTATTTGGCTCTCTCAATGCGCGATGCCAATGCGCTCTATTCC 2168
QY 2297 CACTCCCAACCGGTTTATTAACCGGAAATGTTGGGATCAACGCAAGGCGAAATTTGA 2356
DB 2169 CACCCCTCAGCCGCTTATTAACCGTGAATGTTTGGACACCATGGGAAAAACAAATTTGA 2228
QY 2357 CACGAGCATCACTTTTGTTCAAAGTCGCTATGAATAATGCGTGAAGAAAGCTGGG 2416
DB 2229 CACCAATATCACTTTGTGTTCCAAAGCGCTTCAAGGCGGATCAAGAGAACTAGG 2288
QY 2417 CTTAGAGCGCCAGTCTTACCGGTCAAAACTGCGGTAAACATCAACAAAGAAAGCTTCAA 2476
DB 2289 GCTAGATCGCGCGCACCGCCAGTGAAAACTGTGCAATATCACTAAAAAGGACCTCAA 2348
QY 2477 GTTCAACGCAAAACGGCAAAAAATCACCGTCGATCCGAAAAACCTTCGAGGCTCTTTGTAGA 2536
DB 2349 ATTCAACGATGTGACCGCACATATTGATGTCAACCTGAAACCTTAAAGGTGAAGTGA 2408
QY 2537 TGGCAAACTCTGCACTCTTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTT 2596
DB 2409 TGGCAAGAGTAACTCTTAAAGCAGGAGATGAATTTGAGCTTAGCGCAACTTTTATATTT 2468
QY 2597 CTTCTAGG 2604
|||||

CC having but does not have Helicobacter infection or has a Helicobacter
 CC infection. The attenuated Salmonella vector further comprises an htrA or
 CC ntrB promoter. The vector can be used in inducing an immune response
 CC against Helicobacter in a mammal. The vector can be used to treat
 CC Helicobacter infection. The vector and the method can be used to treat
 CC Helicobacter induced gastroduodenal diseases, including acute, chronic or
 CC atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal
 CC ulcers), and gastric carcinoma. Attenuated Salmonella vector pHUR3 was
 CC constructed expressing UreA and UreB and including an E. coli AmpR gene.
 CC The present sequence represents the pHUR3 plasmid. NOTE: pHUR3 is
 CC represented by both AB55112 and AB55130, the peptides expressed by
 CC pHUR3 are shown in the sequence listing to be split between AB55112 and
 CC AB55130, yet Figure 4 shows all the peptides being expressed by the one
 CC sequence, therefore all encoded peptides are cross-referenced to both
 CC sequences.

XX Sequence 4824 BP; 1389 A; 1090 C; 1146 G; 1199 T; 0 U; 0 Other;

Query Match 34.9%; Score 1005.8; DB 5; Length 4824;
 Best Local Similarity 65.8%; Pred. No. 2.8e-200;
 Matches 1640; Conservative 0; Mismatches 812; Indels 39; Gaps 10;

QY	150	AACGAAATAGGACTATAATCCCATTCCTTTAAATTTTAAACAAG-GAGTAATAGGTG	208
DB	29	AAATGAATCTGACGTACACAGCAATTTAGATATTAAATCATCCACAGGAGAGATCTCCATG	88
QY	209	AAACTCACACCAAGAGCAAGAAAGTTCTTGTATTATATATGCGGCGAAGTGGCTAGA	268
DB	89	AAACTCACACCAAGAGTTAGATAGTTGATGCTCCACTACGCTGGAGAAATGGCTAAA	148
QY	269	AAGCGCAAGACAGAGGGTTAAAGCTCAACCAACCGAAGCCATTCCTTATCATTTAGTGCC	328
DB	149	AAACGCCAAAGAAAGGCATTAAGCTTAATATGTAGAGCAGTAGCTTTGATTTAGTGCC	208
QY	329	CATATTATGACGAGCGCGGTGGAAAAACCGTTGCCAGCTTATGAGAGAGTGC	388
DB	209	CATATTATGGAAGACGAGAGCTGGTAAAAAGACTGCGGTGAATTTGATGCAAGAGGG	268
QY	389	ATGCACATTTTGAAGAAAGATGAAGTAAATGCGCGGGTGGGTAAATATGTTCCCATCTA	448
DB	269	CGCACTCTTTTAAACCGATGATGTGATGCGGTGGCAAGCATGATCATGAGTG	328
QY	449	GGTGTAGAACCACTTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG	508
DB	329	GATATTGAAGCGATGTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG	388
QY	509	CCAGATGAGCACTTCAAGCGCGGGAAGTGAATTTGTTGCGATTAAGACATCGAGCTC	568
DB	389	GCCAAATGGTAAATTTAGTTCTCTGGTGGTTG---TTCTTTAAAAAATGAAGACATCACTATC	445
QY	569	AATGCAGGCAAGAGTAACCGAACTTCAGGTGTTACTAATGAAGGCGCTAAATCTCTTGAT	628
DB	446	AACGAAGCAAAAAGCGGTGAGGTGAAGTTTAAATGTTGGCAGACAGCGGTTCAA	505
QY	629	GTGGGTAGCCATTTCCACTTTCTTGAAGCTAACAGGCACTAAAATTCGATCGTGA AAAA	688
DB	506	ATCGGCTCACACTTTCCATTTCTTGAAGTGAATAGATGCTAGACTTTGACAGAGAAAA	565
QY	689	GCCTATGGCAACCGCTAGATATTCCCTCTCGGCAACAGCTACGATTTGGGCGAGGACAA	748
DB	566	ACTTTGCGTAAACGTTAGACATTTGCGAGCGGACAGCGGTGAAGATTTGAGCTGCGGAA	625
QY	749	ACCCGCAAGTGCAATGATTTCTCTGTCGAGTAAAGAGTGAATTTGGCATGAACGGG	808
DB	626	GA AAAAATCCGTAGATATTGATGACATTTGGCGGTAAACAGAGAAATCTTTGGATTTAACGCA	685
QY	809	CTTGTGATTAACATCGCGG---ATGAAGCGCA-----TAAACA-----TAAAGCG	850
DB	686	TTGGTTGATAGCAAGCAGACAAACGAAGCAAAAAATTTGCTTTACACAGACTTAAGAG	745
QY	851	C-----TTGACAAAGCGG-----AA-----ATCTCAGGATTT---ATCAAGTAAAGGAGACTCCC	896
DB	746	CGTGGTTTTCATGGCGCTAAAAAGCGATGACAACTATGTAAAAAACAATTTAAGGAGTGAAGAA	805

QY	897	ATGAAA---ATGAAAAACAAGAAATATGTAAATACCTACGGACCCCAACCAAGGCGATAAA	953
DB	806	ATGAAAAAGATTAGCAGAAAAAGAAATATGTTCTATGTATGGTCTCTACTACAGGCGATAAA	865
QY	954	GTGCGCTAGGAGATACCGATCTTTGGGCAAGAGTAGACATGACATATACCACTATACCTATGCC	1013
DB	866	GTGAGATTGGGCGATACAGACTTGATCGCTGAAGTAGAACATGACATACCACTATTTATGCC	925
QY	1014	GAAGAACTTAAATTTGGGCGCGGTAAAACTATCCGTGAGGGGTATGGGTGAGCAATAGC	1073
DB	926	GAAGAGCTTAAATTCGGTGGCGGTAAACCCCTAAGAGAAGGCATGAGCCAATCTAACAC	985
QY	1074	CCTGATGAAAAACACCTAGATTATGTCATCACTAAACGCGATGATTATCGACTACACCGGG	1133
DB	986	CCTAGCAAGAAGAGAGTTGGATTTAATTTATCACTAACCGCTTTAATCGTGGATTACACCGGT	1045
QY	1134	ATTTTAAAGCGCATTTGGGATTTAAACCGCAAAATCCATGGCATTTGGCAGGACGGA	1193
DB	1046	ATTTTAAAGCGGATTTTGGTATTAAAGATGGCAAAATCGCTGGCATTTGGTAAAGCGGT	1105
QY	1194	AACAGGACATGCAAGATGGCGTAAGCCCTCATATGTTGCTGGGTGTGGGCAACAGAGCA	1253
DB	1106	AACAAGACATGCAAGATGGCGTTAAACCAATCTTAGGCTAGGTCTCTGCTACTGAAGCC	1165
QY	1254	CTAGCAGGGGAAGGTATGATTTATACCGCTGGGGAAATCGATTTCACACCCACTTTCCTT	1313
DB	1166	TTAGCGGTGAAGGTTTGATCGTAACGGCTGGTGGTATTGACACACATCCACTTCACTTCAVT	1225
QY	1314	TCTCCACAACAATTCCTACCGCTAGCCATGAGGCTTACCAACCATGTTTGGAGGCGGC	1373
DB	1226	TCACCCCAACAATCCCTACAGCTTTTGCAGCGGTGTAACCAACCATGATTTGGTGGGA	1285
QY	1374	ACAGTCTCTGTAGATGGCAAGATCGCACTACTATCACTCCGGGCAAAATGGAATTTGCAC	1433
DB	1286	ACCGTCTCTGTATGGCACTAATCGGACTACTATCACTCCAGGCGAGAGAAATTTAAAA	1345
QY	1434	CGCATGTTGCGCGCAGCAGAGAGATTTCTATGAATTTGGGCTTTTGGGCAAGCAAT	1493
DB	1346	TGGATGCTCAGAGCGCTGAAGAAATTTCTATGAATTTAGTGTCTTGGCTAAAGGTAAC	1405
QY	1494	AGCTCTAGCAAAAACAATTTGTAGAAAGAGTGAAGCGGGCGCGATTTGGTTTAAATTG	1553
DB	1406	GCTTTTAAACGATCGAGCTTAGCCGATCAAAATTTGAAGCGGTGCGATTTGGCAATT	1465
QY	1554	CATGAAGACTGGGCGACCAACCAAGTGCATCGATCACTGCTTGAGCGTGGCAGATGAA	1613
DB	1466	CACGAAGACTGGGCGACCACTCTTCTGCAATCAATCATGCGTTAGATGTTGCGGACAAA	1525
QY	1614	TACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCGAGTTATGTAGATGAC	1673
DB	1526	TACGATGTGCAAGTCCGTATCGCCACAGACACTTTTGAATGAAGCGGTTGTGTAGAGAC	1585
QY	1674	ACCTTAAATGCAATGAACCGGCGCGCATTCATGCTACCAATTTGAGGAGCGGTTGGA	1733
DB	1586	ACTATGGCTGCTATTGCTGGAAGCACTATGACACTTTCCACATGAGGCGCTGCGGCG	1645
QY	1734	GGACACTCACCTGATGTTATCACCATGGCAGGAGCTCAATATTTCTACCTCTCCACC	1793
DB	1646	GGACAGCTCTCTGATTTAATTAAGTAGCCGCTGAACCAACATTTCTCCGCTTCACT	1705
QY	1794	ACCCCACTATTTCCTTATACCAATTAATACGGTTGCAAGAACACTTAGACATGCTCATGACA	1853
DB	1706	AAACCCCACTATCTCTTACCGTGAATACAGAAGCAGACCATGACATGCTTATGGTG	1765
QY	1854	TGCCACCACTAGACAAACGCAATCCGCGAGGATTTTCAATTTTCTCAAGCGGTATCCGC	1913
DB	1766	TGCCACCACTTGGATTAAGAGCAATTAAGAGAGATTTTCACTGCTGCTGATTTCAAGGATCCGC	1825
QY	1914	CCCGCTCTATCGCGCTGAAGATGTCTCCATGATATGGTGTGATCGGATGACAGC	1973
DB	1826	CCTCAAAACCATTTGCGGCTGAAGACACTTTTGCATGATCGGGATTTTCTCAATCACACAGT	1885

Qy	1974	TCGAGTTCGCNAGCAATGGGCGCTGACGGCAAGTGTCTCTCGAACTTGGCGAGACTGCG	2033
Db	1886	TCTGACTCTCAAGCGATTGGGCGGTGTGGGTGAAGTTATATCATAGAACTTGGCAACAGCT	1945
Qy	2034	GATAAGAAATAAAAAGAAATTTGGTAAAGCTTCCTCGAAGATGGCAAGATAACGATAATTC	2093
Db	1946	GACAAAACCAAGAAAGAAATTTGGCGCTTCGAAGAGAGAAAAGCGATACGCAACTTC	2005
Qy	2094	CGCAATTAAGCGCTACATCTCCAAATACACTATCAAGCCCGCTTTTGACCCACGCGGTGAGC	2153
Db	2006	AGATCAAAACGCTACTTGTCTTAAATACACCATTAACCCAGCGATCGTCTCATGGATTAGC	2065
Qy	2154	GAGTATATCGCTCTGTGGAGAGGGCAAGATCGCCGACTTGGTGGTGTGGATCTCTGCC	2213
Db	2066	GAGTATGTAGGTTTCAGTAGAAGTGGGCAAGTGGCTGGACTTGGTATTTGTGGAGTCCAGCA	2125
Qy	2214	TTTTTTGGCGTAAACCACCAAAATCGTGATCAAAAGGGGGTATGTGTGTCTTCTCGAAATG	2273
Db	2126	TTCTTTGGCGTGAACCCCAACATGATCATCAAGGCGGATTCATTGGTTAAGCCAAATG	2185
Qy	2274	GGCGATTCTAACGGGTCTGTGCCACTCCGCCAACCGGTTTATTTACCGCGAAATGTTTGGG	2333
Db	2186	GGCGATGCGAACGCTTCTATCCCTTACCCCAACAACCGGTTTATTTACAGAGAAATGTTTCGT	2245
Qy	2334	CATCACGGCAAGCGGAAATTTTGACACGAGCATCACTTTTGTGTTTCCAAAGTCGCTATGAA	2393
Db	2246	CATCATGTTAAGCTAAATAGATGCAACATCACTTTTGTGTCTCAAGCGGCTTATGAC	2305
Qy	2394	AATGGCGTGAAGAAAGCTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAACTGCCGT	2453
Db	2306	AAAGGCATTAAAGAAATTTAGGACTTTGAAAGACAAGTGTTCGGCTTAAAAAATTCGACA	2365
Qy	2454	AACATCACGAAGAAAGACTTCAAGTTTCAACGACAAAAACGGCAAAATACCGTCGATCCG	2513
Db	2366	AATATCTATAAAAAAGACATGCAATTTCAACGACACTACCGCTCACAATGAAGTCAATCCT	2425
Qy	2514	AAAACTTTCAGGCTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTG	2573
Db	2426	GAACCTTACCATGTGTTCTGGTGGATGGCAAGAAGTAACTTCTTAACACCGCAATAAAGTG	2485
Qy	2574	CCTCTAGCCAGCGCTACACTTTCTTCTAGG	2604
Db	2486	AGCTTGGCGCAACTCTTTTAGCATTTTTCTAGG	2516

RESULT 12	
AEB55112	
ID	AEB55112 standard; DNA; 4824 BP.
XX	
AC	AEB55112;
XX	
DT	22-SEP-2005 (first entry)
XX	
XX	Salmonella vector pHUR3 expressing helicobacter antigens, SEQ ID 1.
DE	
XX	
XX	Circular; pHUR3; ds; vaccine; helicobacter pylori infection; urease;
KW	immune stimulation; antibacterial; Immunostimulant; antiulcer;
KW	cytostatic; antiinflammatory; gastrointestinal-gen.; gastritis;
KW	peptic ulcer; carcinoma.
XX	
XX	Helicobacter pylori.
OS	Salmonella typhimurium.
OS	Escherichia coli.
OS	Synthetic.
XX	
XX	WO200132014-A2.
PN	
XX	
PD	10-MAY-2001.
XX	
XX	
PF	01-NOV-2000; 2000WO-US030191.
XX	
PR	01-NOV-1999; 99US-00431705.
XX	
XX	

(ORAV-) ORAVAX INC.

Kleanthous H, Londono-Arcila P, Freeman D;

WPI: 2001-343379/36

P-FSDB; ABB55113, ABB55114, ABB55115, ABB55116, ABB55117, ABB55118, ABB55119, ABB55120, ABB55121, ABB55122, ABB55123, ABB55124, ABB55125, ABB55126, ABB55127, ABB55128, ABB55129, ABB55131, ABB55132, ABB55133, ABB55134, ABB55135, ABB55136, ABB55137, ABB55138, ABB55139, ABB55140, ABB55141.

Inducing an immune response against Helicobacter in mammals, useful for treating Helicobacter induced gastroduodenal diseases.

Disclosure; SEQ ID NO 1; 63pp; English.

The invention relates to inducing an immune response against Helicobacter in a mammal, comprising mucosally administering to the mammal an attenuated Salmonella vector containing a nucleic acid molecule encoding a Helicobacter antigen, and parenterally administering to the mammal a Helicobacter antigen. Also included is an attenuated Salmonella vector comprising a nucleic acid molecule encoding a Helicobacter antigen. The Helicobacter antigen is a urease, a urease subunit, or its immunogenic fragment (encoded by the ureA and ureB genes). The mammal is at risk of having but does not have Helicobacter infection or has a Helicobacter infection. The attenuated Salmonella vector further comprises an htrA or nifB promoter. The vector can be used in inducing an immune response against Helicobacter in a mammal. The vector can be used to treat Helicobacter infection. The vector and the method can be used to treat Helicobacter induced gastroduodenal diseases, including acute, chronic or atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal ulcers), and gastric carcinoma. Attenuated Salmonella vector pHUR3 was constructed expressing UreA and UreB and including an E. coli Amp^r gene. The present sequence represents the pHUR3 plasmid. NOTE: pHUR3 is represented by both ABB55112 and ABB55130, the peptides expressed by pHUR3 are shown in the sequence listing to be split between ABB55112 and ABB55130, yet Figure 4 shows all the peptides being expressed by the one sequence, therefore all encoded peptides are cross-referenced to both sequences.

Seq Sequence 4824 BP; 1389 A; 1090 C; 1146 G; 1199 T; 0 U; 0 Other;

Query Match 34.9%; Score 1005.8; DB 5; Length 4824;

Best Local Similarity 65.8%; Pred. No. 2.8e-200;

Matches 1640; Conservative 0; Mismatches 812; Indels 39; Gaps 10;

Qy 150 AAGCAATTAGGACTATATCCATTGCCCTTTAAATTTTAAACACAAG-GAGTAATAGGTG 208

Db 29 AAATGAATCTGAGTACACAGCAATTTAGATATTAATCATCATCAGAGAGATCTCCAT 88

Qy 209 AAACACACCCAAAGCAGCAAGAAAGTCTTCTGTATATATTATGCGGCGAAGTGGCTAG 268

Db 89 AAACACACCCAAAGAGTTAGATAGTGTGCTCCATCGCTGGAGAAATTTGGCTAA 148

Qy 269 AAGCGCAAGCAGAGGCTTTAAAGCTCAACCAACCGAAGCCATTCCTTACATAGTGCC 328

Db 149 AAAGCAAGAAAGAGGCATTAGCTTAATGTAGTAGAGCAGTACCTTTGATTAGTGCC 208

Qy 329 CATATTATGCAAGACGGCGCGTGGAAAAAACCCTTTGCCAGCTTTATGGAAGAGTGC 388

Db 209 CATATTATGCAAGACGAGAGCTGCTTAAAGAGCTGCGCTGAATGTATGCAAGAAGG 268

Qy 389 ATGCATCTTTTGAAGAAAGATGAGTAAGTATGCGCGGCTGGTAAATATGCTTCCGATCTA 448

Db 269 CGCAGCTCTTTTAAACCAAGATGATGTGATGGATGGCGTGGCAAGCATGATTCATGAAGTG 328

Qy 449 GGTGTAGAAGCCACCTTTCTCTGATGTGTACAAACTTGTAACTGTGAATTTGGCCCATCGAA 508

Db 329 GGTATTGAAGCATGTTTCTGTATGGGATTAACCTGTAACTGTAACTGTAACTGATCCCTATTGAG 388

Qy 509 CCAGATGAGCATTCAAAGCGGCGAAGTGAAATTTGGTTGGGATAAAGACATCGAGCTC 568

Db 389 GCCAAATGGTAAATAGTCTCTGTGATGTTGTTCTTAAATAAATGAACACATCACTATC 445

QY	569	AATGCAAGGCAAAAGAAAGTAAACCGAATCTGAGGTTTACTAAATGAAGGGCCTAAATCTTTGCAT	628
Db	446	AAGCAAGGCAAAAGAAAGCGGTAGCGTCAAAGTTAAAAATGTTGGCGACAGACCGGTTCAA	505
QY	629	GTGGGTAGCCATTTCCACTTCTTTGAAGCTAAACAAGGCACCTAAAAATTCGATCGTGA AAA	688
Db	506	ATCGGCTCACATTTCCATTTCTTTTGAAGTGAATAGATGCTCTAGACTTTTGACAGAGAAAA	565
QY	689	GCCTATGGCAAAACGCGCTAGATATTCCTCTCTGGCAACAACGCTACGCTATGGGGCAGACAA	748
Db	566	ACTTTCCGTAAAACGCTTAGACATTTGCGAGCGGCACAGCGTAAAGTTTGAGCCTGCGGAA	625
QY	749	ACCGCAAAAGTCAGTTGATTCCTCTTTGGTGGCAGTAAAAAGTGAATTTGGCATGAACGGG	808
Db	626	GA AAAATCCCGTAGAATTTGATTTGACATTTGGCGGTAAACAGAGAATCTTTTGGATTTAAACGCA	685
QY	809	CTTGTGAATAAACATCCGG---ATGAAACGCCA-----TAAACA-----TAAAGCG	850
Db	686	TTGGTTGATAGACAAGCAGACAAAGAAAGCAAAAAAATTCGTTTACAGAGCTAAAGAG	745
QY	851	C-----TTGACAAGCG-----AA-----ATCTCACGGATTT---ATCAAGTAGAGGAGACTCCC	896
Db	746	CGTGGTTTTTCATCGCGCTAAAGCGATGACAACTATGTAAAAACAATTAAGGAGTAAGAA	805
QY	897	ATGAAA---ATGAAAAAACAAGATATGTAAATACCTACGGACCCACCAAGCGGATAAA	953
Db	806	ATGAAAAGAAGATTAGCAGAAAAGAATATGTTTCTATGTATGGTCTCTACTACAGCGCATAA	865
QY	954	GTGCGCTTAGGAGATACCGATCTTTGGCGCAGAAGTAGAACATCAGCTATATCCACTATGGC	1013
Db	866	GTGAGATTGGCGCATACAGACTTGATCGCTGAAGTAGAACATGACTACACCATTTATGGC	925
QY	1014	GAAGAACTTAAATTTTGGCGCGGTAAACATATCCGTGAGGGTATGGGTACAGACCAATAGC	1073
Db	926	GAAGAGCTTAAATTTCCGTGGCGGTAAAAACCTTAAGAGAAGGCATGAGCCAACTCTAACAAAC	985
QY	1074	CCTGATGAAAAACACCCTAGATTTTAGTTCATCACTAAACGCGATGATTAATCGACTACACGGG	1133
Db	986	CCTAGCAAAAGAGATTGGATTTAAATTTATCAGCTAAACGCTTTAATCGTGGATTACACCGGT	1045
QY	1134	ATTTTACAAAGCCACATTTGGGATTTAAAAACGGCAAAATCCATGGCATTTGCAAGGCAGGA	1193
Db	1046	ATTTTTAAAGCGGATTTGGTATTTAAAGATGGCAAAATCGCTGGCATTTGGTAAAGCGGT	1105
QY	1194	AACAAGACATGCAAGATGCGTAAGCCCTCATATGGTCTGGGTGTGGGCACAGAAAGCA	1253
Db	1106	AACAAAGACATGCAAGATGCGTTAAAAACAAATCTTTAGCGTAGGTCTCTGCTACTGAAGCC	1165
QY	1254	CTAGCAGGGGAAGGTATGATTTATACGCTGGGGGAATCGAATTCACACACCACTTTCCITT	1313
Db	1166	TTAGCGCGTGAAGGTTTGATTCGTAACGGCTGGTGGTATTTGACACACACATCCCACTTCATT	1225
QY	1314	TCCTCCAAACAATTCCTACGCTCTAGCCAAATGGCGTTACAAACCATGTTTGGAGGCGGC	1373
Db	1226	TCACCCCAACAATCCCTACAGCTTTTGGCAAGCGGTGAACAACCATGATTTGGTGTGGA	1285
QY	1374	ACAGGTCCTGTAGATGCAAGATGCGACTATCATCTCCGGGCAAAATGGAACTTTGCAC	1433
Db	1286	ACCGGTCCTGCTGATGGCACTAATGCGACTACTATCACTCCAGGCAGAGAAATTTAAAA	1345
QY	1434	CGCATGTTGGCGCAGCAGAGAGTATTCCTATGAAATGTGGCTTTTGGGCAAGGCAT	1493
Db	1346	TGGATGCTCAGAGCGGCTGAAGAAATTTCTATGAAATTTAGGTTTCTTGGCTAAAGGTAAAC	1405
QY	1494	AGCTCTAGCAAAAAACAACCTTGTAGAACAAAGTAGAGCGGGCGCGATTTGGTTTAAATTG	1553
Db	1406	GCTTCTAACCGATGCGAGCTTAGCCGATCAATTTGAAGCGGTGGATGGCTTTGCAATT	1465
QY	1554	CATGAAGACTGGGGCAACAACCAAGTGGCATCGATCACTGTTGAGCGTGGCAGATGAA	1613
Db	1466	CACGAAGACTGGGGCACCACTCTCTTCGCAATCAATCATCGCTTAGATGTTGGGACAAA	1525

RESULT 13
AAQ12485
ID AAQ1
XX

AAQ12485
ID AAQ12485 standard; DNA: 2767 BP.

XX

1614	QY	TACGATGTCAAAGTTGTTATCACACCGATACAGTCAATGAGGCAGGTTATGTAGATGAC	1673
1526	DB	TACGATGTGCAAGTCGCTATCGCCACAGACACTTTTGAATGAAGCCGGTTGTGTAGAAGAC	1585
1674	QY	ACCTAAATGTCAATGAACGGCGCGCCATCCATGCTTACCACTATGAGGAGCGGGTGGGA	1733
1586	DB	ACTATGGCTGTATTGCTGGAAGCACTATGCACTTTTCCACTGAAGCGCTGSGCGG	1645
1734	QY	GGACACTCACTGATGTTATCACCATGGCAGGAGCTCAATATTCTTACCTCTCTCOACC	1793
1646	DB	GGACACGCTCCTGATATTATAAGTAGCCGCTGAACACAACTTCTCCGCTTCCACT	1705
1794	QY	AGCCCCACTATTTCCCTATACCAATTANTAGCGTTGCAGAACCTTACACATGCTCATGACA	1853
1706	DB	AACCCCCACCATCCCTTTACCGTGAATACAGAAGCAGAGCATGTGACATGCTTATGGTG	1765
1854	QY	TGCCACCACTAGACAAACGGCATCCCGGAGGATTTACAATTTTCTCAAAGCCGTATCCCG	1913
1766	DB	TGCCACCACCTTGGATTAAGAGCATTAAGAAAGATGTTTCAGTTCCGCTGATTAAGGATCCCG	1825
1914	QY	CCGGGCTCTATCGCGCTGAAGATGTCTCCATGATATGGGTGTGATTCGGATGACAAAGC	1973
1826	DB	CCTCAAACCAATTCGGCTGAAGACACTTTTGCTATGACATGGGGATTTTCTCAATCACCAGT	1885
1974	QY	TCGGATTCCCAAGCAATGGGCGTGCAGCGGAAGTCATTCCTCGAACTTTGGCAGACTGCG	2033
1886	DB	CTGACTCTCAAGCGATGGCGGTGTGGGTGAAGTTATCACTAGAACTTTGGCAACAGCT	1945
2034	QY	GATAAGAATAAAAAAAATTTGGTAAAGCTTCCTGAAAGATGGCAAGATAACGATAAATTC	2093
1946	DB	GACAAAAACAAGAAAGAAATTTGSCCGCTTTGAAAGAAAGAAAGGCGATTAACGACAATTC	2005
2094	QY	CGCATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGGCGTGAGC	2153
2006	DB	AGGATCAAAACGCTACTGTGTCTAAATACACCATTAACCCAGCGATCGCTCATGGGATTAGC	2065
2154	QY	GAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGCACTTTGGTGTGTGGAATCTCTGCC	2213
2066	DB	GAGTATGTAGTTTCAGTAGAGTGGGCAAGTGCGCTGACTTTGGTATTTGGAGTCCAGCA	2125
2214	QY	TTTTTTTGGCGTAAACCCAAAATCTGTGATCAAGGCGGTAATGGTGGTCTTCTCTGAAATG	2273
2126	DB	TTCTTTTGGCGTGAACCCCAACATGATCATCAAGGCGGATTCATTCGTTAAGCCAAATG	2185
2274	QY	GGGATTTCTAACCGTCTGTGCCACTCTCCCAACCGGTTTATTACCGCGNAATGTTTGGG	2333
2186	DB	GGCGATGCGAAGCGTCTATCCCTACCCCAACAACCGGTTTATTACAGAGAAATGTTTCGT	2245
2334	QY	CATCAGCGCAAGCGGAAATTTTGACACGAGCATCACTTTTGTTCCAAAGTCGCTATGAA	2393
2246	DB	CATCATGTTAAGCTAAATACGATGCAACATCACTTTTGTGTCTCAGCGGCTTATGAC	2305
2394	QY	AATGGCGTGAAGAAAGCTGGCTTTAGAGCGCCAAAGTTCTTACCGGTCAAAAACTGCGGT	2453
2306	DB	AAAGGCAATTAAGAGAAATTAAGGACTTGAAGACAAGTGTGTGCGGTAAAAAAATTCAGA	2365
2454	QY	AACATCACAAGNAGACTTCAAGTTCAAGCAAAAACGGCAAAATCACCGTCTGATCCG	2513
2366	DB	AATATCACTAAAAAAGACATGCAATTTCAACGACACTACCGCTCACATTTGAAGTCAATCCT	2425
2514	QY	AAAACTTCGAGTCTTTGTAGATGGCAAACTCTGCAACCTCTAAACCGCACTCGCAAGTG	2573
2426	DB	GAAACTTACCATGTGTGTGTGGATGGCAAGAAGTAACCTCTAAACCGCAATTAAGTG	2485
2574	QY	CCTCTAGCCAGCGCTACACTTTCTCTAGG	2604
2486	DB	AGCTTGGCGCACTCTTTAGCATTTTCTAGG	2516

AC AAQ12485;
 XX
 DT 27-AUG-2003 (revised)
 DT 23-SEP-1991 (first entry)
 XX
 XX DNA encoding A and B subunits of *H. pylori* urease.
 XX
 KW Gastritis; peptic ulceration; duodenitis; helicobacter; campylobacter; ss.
 XX
 OS Helicobacter pylori.
 XX
 FH Key Location/Qualifiers
 FT CDS 64..717
 FT /*tag= a
 FT /label= subunit A
 FT /note= "66 kD"
 FT CDS 721..2400
 FT /*tag= b
 FT /label= subunit B
 FT /note= "31 kD"
 XX
 PN WO9109049-A.
 XX
 XX 27-JUN-1991.
 PD
 XX 19-DEC-1989; 89GB-00028625.
 PF
 XX 19-DEC-1989; 89GB-00028625.
 PR
 XX (THRE-) 3I RES EXPL LTD.
 XX
 PA Tabaqchal IS, Clayton CL, Wren BW, Kleanthous H;
 PI
 XX WPI; 1991-208084/28.
 DR
 DR P-PSDB; AAR12515, AAR13550.
 XX
 XX Oligo:nucleotide(s) specific for *Helicobacter pylori* - used as probes and
 PT primers to detect *H. pylori* infection, in diagnosis of gastritis, and
 PT duodenal and peptic ulceration.
 PT
 PS Disclosure; Page 16; 28pp; English.
 XX
 XX The DNA is a 2.7 kb TaqI fragment encoding the A and B subunits of *H.*
 CC *pylori* (previously *C. pylori*) urease, i.e. the 66 and 31 kD antigens.
 CC From the sequence, probes and primers can be designed for the
 CC amplification (by PCR) of the gene, to produce a prods. common to all *H.*
 CC *pylori* strains so far tested and which do not occur in other ureases and
 CC can therefore be used as a specific indication of the presence of *H.*
 CC *pylori*. *H. pylori* (previously *C. pylori*) is strongly implicated in the
 CC pathogenesis of gastritis and duodenal and peptic ulceration in man. The
 CC primers/probes can be used for the detection of *H. pylori* DNA in gastric
 CC mucosa, saliva or faecal samples to provide an early diagnosis of
 CC infection. See also AAQ12434-Q12486. (Updated on 27-AUG-2003 to correct
 CC OS field.)
 XX
 XX Sequence 2767 BP; 867 A; 550 C; 635 G; 715 T; 0 U; 0 Other;
 SQ
 Query Match 34.98; Score 1004.8; DB 2; Length 2767;
 Best Local Similarity 66.28; Pred. No. 3.8e-200;
 Matches 1627; Conservative 0; Mismatches 792; Indels 39; Gaps 10;
 QY 182 AAAATTTAACACAGGAGTAAAGTGAAAACCTACACCCCAAGAGCAAGAAAGTTCTTG 241
 DB 40 ACACCTTAAAGATAGGATGATGATGAACTACCCCAAGAGTTAGCAAGTTGATG 99
 QY 242 TTATATATCGGGCGGAGTGGCTAGAAAGCGCAAGCAAGCGGCTTTAAAGCTCAACAA 301
 DB 100 CTCCTACTGCTGAGAAATAGCTTAAAGAAAGCAAGAAAGGCAATTAAGCTTAACTAT 159
 QY 302 CCGAGACCAATTCCTTACATTAGTGCCCATATTATGCAAGAGCGCGCTGGAAGAAAA 361
 DB 160 GTGGAAGCGGTACGTTTGTATTAGTGCCCATATTATGGAAGAGCGAGACGCTGTTAAAG 219

QY 362 ACCGTTGCCAGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCC 421
 DB 220 ACTGCGGCTGAATGATGCAAGAGCGGCACCTCTTTTAAACCGGATGATGATGATGAT 279
 QY 422 GGGGTGGGTAATATGTTCCCGATCTAGGTGTAGAGCCACCTTTCTCTGATGTCGAAA 481
 DB 280 GGGGTGGCAAGCATGATCCATGAAGTGGGTATTGAAGCGATGTTTCTCTGATGGACCAA 339
 QY 482 CTTGTAACTGTGAATTTGGGCCCATCGAACGAGATGAGCACTTCAAGCGGGCGAAGTGAAA 541
 DB 340 CTCGTAAACCGTGCAATACCCCTATTGAGGCAATGTTAAATTTAGTTCTCTGTTGAGTTG 396
 QY 542 TTTGTTGCGATAAAGACATCGAGCTCAATGAGGCAAGAGAGTAAACCGAACCTTTGAGGTT 601
 DB 397 TTTCTAAAAAATGAAGACATCACTATCAACGAAGGCAAGAAAGCCCTTACGCTGAAGATT 456
 QY 602 ACTAATGAAGGCGCTAAATCTTTGATGTGGGTAGCCATTTCCACCTTTCTTTGAAGCTAAC 661
 DB 457 CCCCCTGTTGGGACAGACCGGTTCAATCGGCTCACACTTCCATTTCTTTTGAAGTGAAT 516
 QY 662 AAGGCACTAAAAATTCGATCGTGAAGAAAGCCTATGCAAAAGCGCTAGATATTCCTCTGGC 721
 DB 517 AGATGCTTTAGACTTTGACAGAGAAAAAACTTTCCGTTAAACCGCTTAGACATTTGCGAGCGG 576
 QY 722 AACAGCTACGATTTGGGCGCAGCAACCCGCAAGTGCAGTTGATTTCTCTTGGTGGC 781
 DB 577 ACAGCGGTAAAGTTTGAAGCTTCGCGAAGAAAAATCCGTAGAGTTGATTCACATTTGGCGGT 636
 QY 782 AGTAAAAAAGTGTGGCATGAACGGGCTTTGTGAATTAACATCGCGAT---GAACGCCA- 837
 DB 637 AACAGAGATCTTTTGGATTTAACCGCTTGGTTGATAGCAAGCCGATACGNAAGCAAA 696
 QY 838 -----TAAACA-----TAAAGCGC-----TTGACAAAGCGC-----AA---ATCTCACG 872
 DB 697 AAAATTTGCTTTACAGAGCTAAAGAGCGTGGTTCATGCGCGCTAAAGAGCATGACAAC 756
 QY 873 GATTT---ATCAAGTAAAGAGACTCCCATGAAA---ATGAAAAACAGAAATATGTAAT 926
 DB 757 TATGTAAAAACAAATTAAGGAGTAAAGAAATGAAAAAGATTAGCAGAAAAAGAAATATGCTCT 816
 QY 927 ACCTACGACCCCAAGAGCGATAAAGTGGCTTTAGGAGATACCGATCTTTTGGGCGAGAA 986
 DB 817 ATGTATGCGCTTACTACAGCGGATAAAGTGAATTTGGGCGATACAGACTTGTATCGCTGAA 876
 QY 987 GTAGAAATGACTATACCACTTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTATC 1046
 DB 877 GTAGAAATGACTACACCATTTATGTTGAAGAGCTTAAATTTGGCGCGGTAAAAACCTTA 936
 QY 1047 CGTGAAGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACCCCTAGATTTAGTCACTACT 1106
 DB 937 AGAGAAGGCTAGAGCAATCTAACAAACCTTAGCAAAAGAAAGTGTGATCTAATCATCACT 996
 QY 1107 AACCGGATGATTGCACTACACCGGATTTACAAAGCGGATTTGGGATTTAAAAACGCG 1166
 DB 997 AACGCTTTAATCGTGGATTTACCGCGTATTATTAAGCGGATTTGGTATTAAAGATGGC 1056
 QY 1167 AAAATCCATGCTTTGGCAAGCGAGGAAACAGGACATGCAAGATGGCGCTTAAGCCCTCAT 1226
 DB 1057 AAAATCGCTGCTTTGTTAAAGCGGTAAACAAAGACACGCAAGATGGCGTTAAAAACAT 1116
 QY 1227 ATGTCGTGGGTGTGGGCACAGAGCACTAGCAGGGGAGGTATGATTTATTACCGCTGGG 1286
 DB 1117 CTTAGCGTGGGTCTCTGCTACTTGAAGCCCTTAGCGGTGAAGGTTTGTATTGTAACCTGCTGGT 1176
 QY 1287 GGAATCGATTCACACACCACTTCTTCTTCTCAAAATTTCCCTACCGCTCTAGGCAAT 1346
 DB 1177 GGTATTGACACACATCTCACTTCTTCTCCCGCAACAAATCCCTACAGCTTTTGAAGC 1236
 QY 1347 GGCCTTACAAACCATGTTTGGAGGCGGACAGCTCTCTGTAGATGGCAGCAAGATCGCACTACT 1406
 DB 1237 GGTGTAAACCAATGATTGGTGGCGGAACCTGCGCTGCTGATGGCACTAACGCAACCACT 1296
 QY 1407 ATCACTCCGGGCAAAATGGAACTTGCAACCGCATGTTGCGCGCAGCAGAGAGTATTCTATG 1466

Db 1297 ATCAGTCCAGGTGAAGAAATTTAAATTCATGCTCAGAGCGGTGAAGAAATATTCATATG 1356
Qy 1467 AATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAAACAACCTTTGTAGAACAAAGTA 1526
Db 1357 AACTTGGTTCTTGGCTAAAGGTAAAGCTCTTAACGATGCAAGCTTAGCCGATCAAAAT 1416
Qy 1527 GAAGCGGCGCGATGTTGTTTAAATTCATGAAGACTGGGGCAACAAACCAAGTGGCATC 1586
Db 1417 GAAGCTGGTGGATGTCCTTTAAATTCACGAAGACTGGGGCAACCACTCTCTTCTGCAATC 1476
Qy 1587 GATCAGTCTGCTGAGCGTGGCAGATGATAGATGTCGAAGTTTGTATGCCACACGATACA 1646
Db 1477 AATCATGCGTTAGATGTTGGGGAACAAATAGATGTCAAGTCGCTATCCACACAGACACT 1536
Qy 1647 GTCAATGAGGCGAGTTATGTAGATGACACCCCTAAATGAACGCGGCGGCATCCAT 1706
Db 1537 TTGAATGAGCGGTTGCTGGTGAAGACACTATGCGAGCTATTCGCGGACGACTATGCAC 1596
Qy 1707 GCTTACACATGAGGAGCGGTTGAGGACACTCAGCTGATGTTATCACCATGGCAGGC 1766
Db 1597 ACTTACCACACTGAAGCGCTGGCGGCGGACACGCTCTGATATTAATAAGTGGCGGT 1656
Qy 1767 GAGCTCAATTTCTCAGCTCTCCACACCCCACTATTCCTATACCATTAATACGGTT 1826
Db 1657 GAACACAAACATCTCCCGCTTCCACTAAACCCCACTATCCCTTTACCGTGAATACAGAA 1716
Qy 1827 GCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGAT 1886
Db 1717 GCCGAACATGACATGCTTATGCTGTGCCACCACTTGATTAAGCAATTAAGAAGAT 1776
Qy 1887 TTACAATTTTCTCAAAGCGTATCCGCGCGGCTCTATCGCGCTGAAGATGTCTCCAT 1946
Db 1777 GTCCAGTTGCTGATTCAAGGATTCGCGCTCAACACCATTCGCGCTGAAGACACTTTGCAT 1836
Qy 1947 GATATGGGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGTGCAGGCGAA 2006
Db 1837 GACATGGGGATTTTCTCAATCACTAGTTCTGACTCTCAAGCGATGGCGGTGGGTGAA 1896
Qy 2007 GTGATTCCTCGAATCTGCGAGCTGCGATGAAGATTAAGAAAGATTTGGTAAGCTTCT 2066
Db 1897 GTTATCACTAGACTTGGCAACAGCTGACAAAAATTAAGAAATTTGGCGCTTGAA 1956
Qy 2067 GAAGATGGCAAGATAACGATAATTTCCGATTAAGCGCTACATCTCCAAATACATATC 2126
Db 1957 GAAGAAAAGCGGATACGACAACTTCAGGATCAACGCTACTTGTCTAATAACCAATT 2016
Qy 2127 AACCCCGCTTTGACCCACGCGGTGAGGAGTATATCGGCTCTGTGGAAGGGCAAGATC 2186
Db 2017 AACCCAGCGATCGCTCATGGATTAGCGAGTATGTCGGTCTGTAGAGTGGGCAAGTG 2076
Qy 2187 GCCGACTGGTGGTGGATCCTGCTCTTTTGGCGTAAACCCAAATCGTATCAAA 2246
Db 2077 GCTGACTGTGTTTGTGGAGTCCGCGATCTTGGTGTGAACCCAAACATGATCATCAA 2136
Qy 2247 GGCGGTATGGTGGTCTCTGAAATGGCGATTTCAACGCGTCTGTGCCCACTCCCCAA 2306
Db 2137 GGCGGTTTCACTGCAATTCAGTCAAAATGGGTGATGGGAACGTTCTATCCCTACCCCAA 2196
Qy 2307 CCGGTTTATTACCGCAAAATTTTGGGCATCACGCAAGGCGAAATTTGACACAGCATC 2366
Db 2197 CCAGTTTATTACAGAGAAATGTCGCTCATCATGTTAAAGCTAAATACGATGCAACATC 2256
Qy 2367 ACTTTTGTCCAAAGTCCCTATGAATAATCGCGTGAAGAAAGCTGGGCTTAGAGCGC 2426
Db 2257 ACTTTTGTCTCAAGCGGCTTATGACAAAGGCAATTAAGAAAGATTAAGGCGTTGAAGA 2316
Qy 2427 CAAGTTCTACCGGTCAAAACCTGCGTAAACATCACCAAGAAAGACTTCAAGTTCACAGAC 2486
Db 2317 CAAGTGTTCGCGGTAAATAATTTGAGAAATATACATAAAGAAAGATCAATTCATACGAC 2376
Qy 2487 AAAACGGCAAAATACACGTCGATCCGAAAAACCTTCGAGGTCTTTGTAGATGGCAAACTC 2546

Db 2377 ACTACCGCTCACATTGAAGTCAATTTCTGAACCTTACCATGTGTTCTGTGATGCCAAA-GA 2435
Qy 2547 TGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCGCGCTACACTTTTCTTAGG 2604
Db 2436 AGTAACTCTAAACCCACCAATAAAGTGAAGCTTGGCGCAACTCTTTAGCATTTTCTAGG 2493
RESULT 14
ID ADQ37850 standard; DNA; 1710 BP.
XX ADQ37850;
AC ADQ37850;
XX 07-OCT-2004 (first entry)
DT H. bizzozeronii ureB DNA.
DE Urease; urease gene cluster; urease structural gene;
KW urease accessory gene; ureB; Helicobacter bizzozeronii infection;
KW antibacterial; gene; ds.
XX OS Helicobacter bizzozeronii.
XX Key Location/Qualifiers
FH CDS 1..1710
FT /*tag= a
FT /product= "H. bizzozeronii ureB"
PN US2004142343-A1.
XX 22-JUL-2004.
XX 12-AUG-2003; 2003US-00639273.
XX 16-AUG-2002; 2002US-0404337P.
XX (CHAN/) CHANG Y.
PA (SIMP/) SIMPSON K W.
PA (ZHUJ/) ZHU J.
XX Chang Y, Simpson KW, Zhu J;
XX WPI; 2004-533502/51.
DR P-PSDB; ADQ37851.
XX Novel isolated nucleic acid molecule having urease gene cluster, and
PT conferring on Helicobacter bizzozeronii ability to produce urease, useful
PT as vaccine for preventing disease in mammals infected by H.bizzozeronii.
XX Claim 7; SEQ ID NO 4; 40pp; English.
XX The invention relates to an isolated nucleic acid molecule conferring on
CC Helicobacter bizzozeronii an ability to produce urease, where the nucleic
CC acid molecule is a urease gene cluster comprising at least one urease
CC structural gene and at least one urease accessory gene. The nucleic acid
CC molecule is chosen from ureA, ureB, ureF, ureG, ureH and ureI. The
CC invention also relates to an isolated protein encoded by the nucleic
CC acid, a vaccine for preventing onset of disease in mammals infected by H.
CC bizzozeronii comprising a nucleic acid and a carrier, and an isolated
CC antibody or its binding portion raised against the nucleic acid. The
CC nucleic acids, proteins and antibodies are useful for vaccinating mammals
CC against onset of disease caused by infection of H. bizzozeronii, which
CC involves administering the sequences. The sequences are useful for
CC detecting H. bizzozeronii in a sample of tissue or body fluids which
CC involves providing a nucleic acid as an antigen, providing an antibody,
CC or providing a nucleotide sequence as a probe in a nucleic acid
CC hybridisation assay, contacting the sample with the antigen or the probe,
CC and detecting any reaction which indicates that H. bizzozeronii is
CC present in the sample. This sequence represents H. bizzozeronii ureB DNA.
SQ Sequence 1710 BP; 492 A; 436 C; 386 G; 396 T; 0 U; 0 Other;
Query Match 32.7%; Score 940.8; DB 12; Length 1710;

[illegible]

PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
 XX Mckenzie B, Boyle J, Lew A;
 XX WPI; 2003-156789/15.
 XX Raising an immune response in an animal by administering composition
 PT comprising carrier and antigen bound to targeting moiety which binds to
 PT receptor present in circulatory vessels in Gut Associated Lymphoid
 PT tissue.
 XX Disclosure; Page 27-28; 45pp; English.
 XX This sequence encodes H. felis urease B. This sequence was used in the
 CC method of the invention for raising an immune response in an animal. The
 CC method comprises administering to the animal a composition comprising a
 CC carrier and an antigen bound to a targeting moiety which binds to at
 CC least one receptor present in circulatory vessels in Gut Associated
 CC Lymphoid Tissue (GALT). The method is useful for raising an immune
 CC response in an animal against antigens from Salmonella, Cholera,
 CC Helicobacter pylori, HIV, Candida, P. gingivalis, gut parasites, gut
 CC associated toxins, gut hormones, gut hormone receptors or gut associated
 CC cancers. The method is useful for raising both mucosal and systemic
 CC immunity against any antigen used in the composition
 XX
 SQ Sequence 1719 BP; 490 A; 411 C; 421 G; 397 T; 0 U; 0 Other;
 Query Match 31.3%; Score 902.4; DB 8; Length 1719;
 Best Local Similarity 70.6%; Pred. No. 9.1e-179;
 Matches 1203; Conservative 0; Mismatches 501; Indels 0; Gaps 0;
 900 AAAATGAAAAAACAAGATATGTAATACCTACCGACCCCAAGGGGATGAAGTGGCG 959
 15 AAGATTTACGAAAGAAATATGTTCTATGATGTCCTCCACTACCGGGATCGTGTAGA 74
 960 TTAGAGATACCGATCTTTGGGAGAGATGAGATGATGATGATGATGATGATGATGAT 1019
 75 CTCGGCGACACTGATTTGATCTTAGAAGTGGAGCATGATGATGATGATGATGATGAT 134
 1020 CTTAAATTTGGCGGGTAAACATCTCGTGGGGTATGGGTGACGAAATAGCCCTGAT 1079
 135 ATCAAAATTTGGGGCGGTAACATCTCGTGGATGGATGATGATGATGATGATGATGAT 194
 1080 GAAACACCCCTAGATTTAGTATCATCACTAACCGGATGATTTATCGATCAACCGGGATTTAC 1139
 195 TCTTATGATTTAGATTTGGTCTCACTAACGCCCTCATTTGGATATACGGGCAATTTAC 254
 1140 AAAGCGGATTTGGATTTAAACCGGCAAAATCCATGGCATTTGGCAAGGAGGAAACAG 1199
 255 AAAGCGGATTTGGATTTAAACCGGCAAAATTTGGCAAGGAGGAGGCAATTAAG 314
 1200 GACATGCAAGATGGGTAAAGCCCTCATATGTCGTGGGTGGGTCAGACAGACACTAGCA 1259
 315 GACATGCAAGATGGGTAAAGCCCTCATATGTCGTGGGTGGGTCAGACAGACACTAGCA 374
 1260 GGGGAAGTATGATTTATACCGTGGGGAATCGATTCACACACCCACTTCTCTTCTCCA 1319
 375 GCTGAGGCTTTGATTTAAACCGCTGGTGGCATCGATACGATATTTCACTTTATCTCTCC 434
 1320 CAACAATTTCCCTACCGCTCTAGCCAAATGGGTTAAACCAATGTTGGAGCGGCAAGGT 1379
 435 CAACAATTTCCCTACTGCTTTTGGCAGCGGGTTACAAACCATGATTTGGAGGAGGCAAG 494
 1380 CTTGATGATGCAAGATTCGACACTACTATCACTCCGGGCAATGGAATTTGGCAGCGCATG 1439
 495 CTTGGATGCAAGATTCGACACTACTATCACTCCGGGCAATGGAATTTGGCAGCGCATG 554
 1440 TTGCGCGCAGCAGAGATTTCTATGATGTTGGCTTTTGGGCAAGGCAATAGCTCT 1499
 555 TTGCTGCGCGCAGCAGATTCGCAATGATCTAGGCTTTTGGCTTAAGGGGATGTGCT 614
 1500 AGCAAAAAACAACCTTTGTAGAAACAAGTAGAAGCGGGCGGATTTGGTTTAAATTTGCATGAA 1559

Db 615 TAGGAACCCCTCTTTAGCGGATCAGATTGAAGAGGGGGGATGTTGTTTAAATCCAGAA 674
 QY 1560 GACTGGGGCACAACCAACCAAGTGGATCGATCACTGCTTGAGCGTGGCAGATGAATACGAT 1619
 Db 675 GACTGGGGAAGCAGACACCTGCGAGCTATTACCACTGCTCAATGTCGCGATGAATACGAT 734
 QY 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGAGTTATGTAGATGACACCCCTA 1679
 Db 735 GTGCAAGTGGCTATCCACACCGATACCTTTAAAGAGGCGGGCTGTGTAGAGACACCCCTA 794
 QY 1680 AATGCAATGAACGGGCGGCCATCCATGCTTACCACTTGAAGGAGGCGGTGGAGGACAC 1739
 Db 795 GAGGCGATTTCCGCGCGGACCACTTCACTTCCACATGAGGGGCTGGGGTGGAGACAC 854
 QY 1740 TCACCTGATTTATCACCATGGCAGCGAGCTCAATATTCTACCCCTCTCCACACACCCCC 1799
 Db 855 GCTCCAGATTTATCAAAATGGCAGGGAATTTAACTTCTACCCGCTCTTACTAACCCTG 914
 QY 1800 ACTATTTCCCTATACCAATTAATAGGTTGAGAGACACTTAGACATGCTCATGACATGCGAC 1859
 Db 915 ACCATTTCTTTTACCACCAAAACACTGAAGCGGAGACATGAGACATGTTAATGGTGTGCGAC 974
 QY 1860 CACCTAGACAAAACGCGATCGCGAGGATTTACAAATTTTCTCAAGCCGTATCCCGCCCGG 1919
 Db 975 CACTTGGATAAAAGTATCAAGGAAGATGTGAGTTTCCGATTCGAGGATTCGCCCCCAA 1034
 QY 1920 TCTATCGCGGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATGACAAGCTCGGAT 1979
 Db 1035 ACTATCGCGGCTGAAGACCAACTCCATGATGATGGGATCTTTTCTATCACCAGCTCGAC 1094
 QY 1980 TCGAAGCAATGGGCGTGCAGCGGAGATGATTTCTCGAATCTTGGCAGACTGGGATAAG 2039
 Db 1095 TCTCAGGCTATGGGCGCGCTAGCGAGGATGATCACACGCACTTTGGCAGACAGACAAA 1154
 QY 2040 AATAAAAAGATTTGGTAAAGCTTCTGGAAGATGGCAAGATTAACGATAATTTCCCGATT 2099
 Db 1155 AAAAAAGAGTTTGGGCGCTTGAAGAGGAAAAAGCGGATTAACGACAACTTCCGATC 1214
 QY 2100 AAGCGCTACATCTCAAAATACACTATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTAT 2159
 Db 1215 AAGCGTACATCTTAATATACACCATCAACCCCGCGATCGCGCATGGATTTCTGACTAT 1274
 QY 2160 ATCGGCTCTGTGGAAGAGGCAAGATGCGCGATTTGGTGGTGTGGAATCTCGCTTTTCT 2219
 Db 1275 GTGGGCTCTGTGGAAGTGGCAAAATACGCGACCTCGTGTCTTGGAGTCCGGCTTTCTTT 1334
 QY 2220 GGGTAAAAACCAATCGTATCAAGCGGATGATGGTGTCTTCTCTGAAATGGGCGAT 2279
 Db 1335 GGCATTAAGCCCAATATGATTTAAGGGCGGATTTATTCGGCTCTCTCAATGGGCGAT 1394
 QY 2280 TCTAAGCGCTCTGTGCCACTCCCAACCGGTTTATTTACCGCAAAATGTTTGGGCAATC 2339
 Db 1395 GCAATGCTCTATTCTCCACCCCTCAGCCGCTATTTACCGTGAATGTTTGGACCAT 1454
 QY 2340 GGAAGGGGAAATTTGACACACGATCATCTTTTGTGTTTCCAAAGTCGCTATGAAATGGC 2399
 Db 1455 GGAAGAAACAAATTCGACACCAATATCACTTTGTGTGTCGCAAGCGGCTTACAAGCGAGG 1514
 QY 2400 GTGAAAGAAAGCTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAAATGCGCGTAAATC 2459
 Db 1515 ATCAAGAGAGAACTAGGGCTAGATCGGTGGTATTCGAGTGAAGAAATCTGTGCAATATC 1574
 QY 2460 ACCAAGAAAGACTTCAAGTTTCAACGACAAACCGGCAAAATCAACCGTCAATCCGAAAC 2519
 Db 1575 ACTAAAGAGAACTCAATTTCAACGATGACCGCATATTTGATGATCAACCTTGAAC 1634
 QY 2520 TTGAGGCTCTTTGTAGATGGCAAACTCTGACCTCTAAACCCCACTCGCAAGTGCCTCTA 2579
 Db 1635 TATAAGTGAAGTGAAGGAGGTAACCTCTAAAGCAGCAGATGAATTTGAGCCTA 1694
 QY 2580 GCCAGCGCTACACTTTCTCTAG 2603
 Db 1695 GCGCACTTTATAATTTGTTCTAG 1718

Search completed: August 10, 2006, 01:43:00
Job time : 1755 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2006, 01:37:34 ; Search time 528 Seconds
(without alignments)

10216.679 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2880.6

Sequence: 1 rgragattttccaractt.....aaaaagtagaacccacagg 2883

Scoring table: IDENTITY NUC

Gapop 3.0 , Gapext 3.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*
1: /EMC_Celerra_SIDS3/prodata/2/ina/1_COMB.seq:*
2: /EMC_Celerra_SIDS3/prodata/2/ina/5_COMB.seq:*
3: /EMC_Celerra_SIDS3/prodata/2/ina/6A_COMB.seq:*
4: /EMC_Celerra_SIDS3/prodata/2/ina/6B_COMB.seq:*
5: /EMC_Celerra_SIDS3/prodata/2/ina/7_COMB.seq:*
6: /EMC_Celerra_SIDS3/prodata/2/ina/8_COMB.seq:*
7: /EMC_Celerra_SIDS3/prodata/2/ina/PCUS_COMB.seq:*
8: /EMC_Celerra_SIDS3/prodata/2/ina/PP_COMB.seq:*
9: /EMC_Celerra_SIDS3/prodata/2/ina/RE_COMB.seq:*
10: /EMC_Celerra_SIDS3/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1021.4	35.5	2735	2	US-08-920-095-1
2	1021.4	35.5	2735	7	PCT-US96-05800-1
3	1008.8	35.0	2619	2	US-08-467-822-19
4	1008.8	35.0	2619	3	US-08-432-697-19
5	1008.8	35.0	2619	3	US-08-466-248-19
6	1005.8	34.9	4824	3	US-09-431-705-1
7	1005.8	34.9	4824	3	US-09-431-705-1
8	667.4	23.2	6131	2	US-07-732-242C-8
9	665.4	23.1	1830121	3	US-09-557-884-1
10	665.4	23.1	1830121	3	US-09-643-990A-1
11	665.4	23.1	1830121	3	US-10-158-865-1
12	648.2	22.5	1710	3	US-09-543-681A-1857
13	617.2	21.4	2400	2	US-08-967-513-1
14	617.2	21.4	2400	2	US-08-687-645B-1
15	612.8	21.3	1878	3	US-09-489-039A-2045
16	610.2	21.2	1716	3	US-08-487-429A-3
17	610.2	21.2	1716	7	PCT-US96-05320A-541
18	593.6	20.6	5966	3	US-08-956-171B-22
19	593.6	20.6	5966	3	US-08-781-986A-22
20	573.4	19.9	1704	3	US-09-328-352-1786
21	569.8	19.8	8729	3	US-09-453-702B-258
22	569.8	19.8	8729	3	US-10-114-170-258
23	569.8	19.8	87563	3	US-09-453-702B-57

24	569.8	19.8	87563	3	US-10-114-170-57	Sequence 57, Appl
25	567.6	19.7	1701	3	US-09-252-991A-10316	Sequence 10316, A
26	563.8	19.6	3164	3	US-09-710-279-3828	Sequence 3828, Ap
27	563.8	19.6	3234	3	US-09-710-279-3853	Sequence 3853, Ap
28	562.8	19.5	1716	3	US-09-710-279-117	Sequence 117, App
29	562.8	19.5	1722	3	US-09-134-001C-2189	Sequence 2189, Ap
30	552	19.2	1653	3	US-09-252-991A-10546	Sequence 10546, A
31	542	18.8	1833	3	US-09-602-777A-13	Sequence 13, Appl
32	518.6	18.0	1625	3	US-09-602-777A-15	Sequence 15, Appl
33	473.8	16.4	1797	3	US-09-601-198-89	Sequence 89, Appl
34	470.6	16.3	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
35	470.6	16.3	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
36	374.2	13.0	3475	3	US-09-710-279-4032	Sequence 4032, Ap
37	356.6	12.4	951	3	US-09-252-991A-10513	Sequence 10513, A
38	211.2	7.3	684	3	US-10-012-819-153	Sequence 153, App
39	208.2	7.2	636	3	US-09-252-991A-10653	Sequence 10653, A
40	173.2	6.0	363	3	US-10-012-819-159	Sequence 159, App
41	106.8	3.7	462	3	US-09-543-681A-1913	Sequence 1913, Ap
42	106.6	3.7	613	3	US-09-533-559-7520	Sequence 7520, Ap
43	100.2	3.5	312	3	US-09-328-352-1761	Sequence 1761, Ap
44	99.6	3.5	360	3	US-09-328-352-1789	Sequence 1789, Ap
45	98.8	3.4	399	3	US-09-134-001C-2212	Sequence 2212, Ap

ALIGNMENTS

RESULT 1

US-08-920-095-1
; Sequence 1, Application US/08920095
; Patent No. 5837240
; GENERAL INFORMATION:
; APPLICANT: Cynthia K. Lee et al.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,095
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,041
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2735 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-920-095-1

Query Match 35.5%; Score 1021.4; DB 2; Length 2735;
Best Local Similarity 66.5%; Pred. No. 6.6e-244;
Matches 1634; Conservative 0; Mismatches 786; Indels 38; Gaps 9;

Qy	182	AAAATTAA	CACAAGGAGTAA	TAGGTGAAA	CTCACC	CAAGAGCA	AGAAAAAGTTCTTG	241
Db	78	ACACTTTA	AGAAATAGGAGA	ATGAGTGA	AACTCACC	CAAAAGAGT	TAGATAAGTTGATG	137
Qy	242	TTATATAT	TATCGGGCGA	AGTGGCTAG	AAAGCGGCA	AGCAGAGG	CGCTTAAAGCTCAACCA	301
Db	138	CTCACAT	CACCTCGAGAA	TATGGCTTAA	AAAAACGCA	AGAAAAAGCA	TTAAGCTTAACTAT	197
Qy	302	CCGAAAGC	CAATGTCTTAC	ATATAGTGC	CCATATAT	TGACGAAG	CGCGCGTGGAAAAAAA	361
Db	198	GTAGAGCA	GTAGCTTTGAT	TAGTGCCTAT	TATGGAAG	AGCGAGAG	CTGGTAAAGAG	257
Qy	362	ACGTTGCC	CAGCTTATG	GAAGAGTGC	ATGCATCT	TTTTTGAAAA	AGAGTAAGTAATGCCCC	421
Db	258	ACTCGGCT	GAAATGATG	CAAGAGGCG	CACCTCTTTT	TAAAAAC	CCAGATGATGATGGAT	317
Qy	422	GGGTGGGT	TAATATGTTT	CCCGATCT	AGGTGTAG	AAAGCCAC	CTTCTGATGGTACGAAA	481
Db	318	GGGTGCGA	CAGCATGAT	CCATGAAG	TGGGTATT	TGAAGCGAT	GTCTTCTGATGGGACTAAA	377
Qy	482	CTTGTAAC	TGTGAATGG	CCCATCG	AACAGATG	GAGCACCT	TCAAAGCGGCGCAAGTCAAA	541
Db	378	CTCGTAAC	CGGTGCA	TACCCCTAT	TGAGGCCA	ATGGTAAAT	TAGTCTCTGGTAGTTG	434
Qy	542	TTTGGTTG	CGATAAAGCA	CATCGAGCT	CAATG	CGAGGCA	AGAAGTAAACCGAATTTAGGTT	601
Db	435	TTCTTAAAA	AAATGAAGACA	TCACTAT	CAACGAAGG	CAAAAAAG	CCGTTAGCGTGAAGATT	494
Qy	602	ACTAATGA	AGGGCTTAAT	CTTGTCAT	TGTGGGTAG	CCATTTCC	NCTTCTTTGAAGCTAAC	661
Db	495	AAAAATTT	TGGCAGAC	AGCCGGTT	TCAAATCG	GCTCACACT	TCCATTTCTTTGAAGTGAAT	554
Qy	662	AAGSCACT	AAAAATTC	GTATCGTGT	AAAAAGCCT	TATGGCAAC	CGCTAGATATTTCCCTCTGGC	721
Db	555	AGATGCC	TAGACTTTG	ACAGAGAAAA	AACTTT	CGGTAAAC	GCTTAGACATTCGAGCGGG	614
Qy	722	AACACGCT	ATCGCATTT	GGGGCAG	CAAAACCG	CAAAAGTGC	AGTTGATTCCTCTTGGTGGC	781
Db	615	ACAGCGGT	TAAGATTTG	AGCTTGG	CGAGAAAA	ATCCGTAG	ATTTGATTTGACATTTGGCGGT	674
Qy	782	AGTAAAAA	AGTGATTTG	GCATGA	ACGGGCTGT	GTGAATAA	CAATTCGCGG---ATGAACGCCA-	837
Db	675	AACAGAGA	ATCTTTTGG	ATTTAA	CGCATTTGGT	TGTATAG	ACAAGCAGACAACGAAGCAAA	734
Qy	838	-----	TAACA----	TAAGCG----	TTGACA	AGCG-----	AA---ATCTCACG	872
Db	735	AAATTGCT	TTACACAG	AGCTTAA	AGACGTGGT	TTTTCAT	GGCGCTAAAGCGGATGAAC	794
Qy	873	GATTT---	ATCAAGTA	AGGAGACT	CCCATG	AAAA---ATG	AAAAAACAAGAAATATGTAAT	926
Db	795	TATGTAAA	ACAAATTA	AGGTAG	TAAGAAAT	TGAAAAGAT	TAGCAGAAAAGAAATATGTTCT	854
Qy	927	ACCTACGA	CCCCAC	CAAGCG	ATAAAGT	TCGCTTAG	GAGATACCGATCTTTGGGCGAGAA	986
Db	855	ATGTATGT	CTCTACT	ACAGCG	CGATAA	AGTGTAG	ATTGGCGGATACAGACTTTGATCGCTGAA	914
Qy	987	GTAGAA	CATGACTAT	ACCACT	TGGCGA	AGAACTTAAAT	TTGGCGGGGTAAACTATC	1046
Db	915	GTAGAA	CATGACT	ACCACT	TTATGG	CGAAGAGCTTAAAT	TTCGGTGGCGGTAAACCCCTA	974
Qy	1047	CGTGAGGT	ATGGGT	CAGAGCA	ATAGCCCT	TGATG	AAAAACCCCTAGATTTTAGTCACT	1106
Db	975	AGAGAGCA	TAGCCAA	TCTAAC	ACCTTAG	CAAAAGAG	AGTTGGATTTAATTTACT	1034
Qy	1107	AACGCGAT	GTATAT	TCGACT	ACCGGAT	TTTACA	AAAGCGCAATTTGGATTTAAAAACGGC	1166
Db	1035	AACGCTTT	TAATCGTGG	ATTAC	CCGGTATTT	TATAAAGCG	GATATTGGTATTTAAAGATGCG	1094
Qy	1167	AAAAATCA	TGGCATTTG	CAAGG	CAGGAAAA	CAAGACAT	GCNAGATGCGGTAAAGCCCTCAT	1226
Db	1095	AAAAATCG	TGGCATTTG	TAAAGG	CGGTAA	CAAAAGCA	TGCAAGATGCGTTTAAAAACAAT	1154

Qy	1227	ATG	CTCTG	TGGGTGTGGG	CA	GA	AA	GC	ACTAG	C	AGG	GG	AA	GGTAT	GA	TAT	TAC	CG	CTGGG	1286
Db	1155	C	T	TAG	CGT	TAG	TCTCTG	TACT	TGA	AG	C	T	T	TAG	CG	CG	T	T	GA	1214
Qy	1287	G	GA	AT	TC	GA	T	CA	AT	T	T	T	T	T	T	T	T	T	T	1346
Db	1215	G	GA	T	T	G	A	C	A	C	A	T	T	T	T	T	T	T	T	1274
Qy	1347	G	G	G	G	T	T	A	A	C	C	A	T	T	T	T	T	T	T	1406
Db	1275	G	G	T	G	T	A	A	C	C	A	T	T	T	T	T	T	T	T	1334
Qy	1407	A	T	C	A	C	T	C	G	G	G	G	A	A	T	G	A	A	T	1466
Db	1335	A	T	C	A	C	T	C	G	G	G	G	A	A	T	T	T	T	T	1394
Qy	1467	A	A	T	G	T	G	G	C	T	T	T	T	T	T	T	T	T	T	1526
Db	1395	A	A	T	T	G	T	T	C	T	T	T	T	T	T	T	T	T	T	1454
Qy	1527	G	A	A	C	G	G	G	C	G	A	T	T	T	T	T	T	T	T	1586
Db	1455	G	A	A	C	G	G	T	G	G	A	T	T	T	T	T	T	T	T	1514
Qy	1587	G	A	T	C	A	C	T	G	C	T	G	G	A	T	G	A	T	T	1646
Db	1515	A	A	T	C	A	T	G	C	T	G	G	A	T	G	A	T	T	T	1574
Qy	1647	G	T	C	A	A	T	G	A	T	G	A	T	G	A	T	G	A	T	1706
Db	1575	T	T	G	A	T	G	A	A	C	C	G	G	T	G	T	A	T	T	1634
Qy	1707	G	C	T	T	A	C	A	C	A	T	T	T	T	T	T	T	T	T	1766
Db	1635	A	C	T	T	T	C	A	C	A	T	T	T	T	T	T	T	T	T	1694
Qy	1767	G	A	G	C	T	A	A	T	T	C	C	T	C	C	A	C	A	T	1826
Db	1695	G	A	A	C	A	A	C	A	T	T	T	T	T	T	T	T	T	T	1754
Qy	1827	G	C	A	A	C	A	T	T	G	A	T	G	A	T	G	A	T	T	1886
Db	1755	G	C	A	G	A	C	A	T	G	A	T	G	A	T	G	A	T	T	1814
Qy	1887	T	T	C	A	A	T	T	T	C	A	A	C	C	G	T	C	T	T	1946
Db	1815	G	T	T	C	A	G	T	T	T	C	A	A	C	C	T	C	A	T	1874
Qy	1947	G	A	T	G	G	T	G	A	T	G	A	T	G	A	T	G	A	T	2006
Db	1875	G	A	C	A	T	G	G	A	T	T	T	T	T	T	T	T	T	T	1934
Qy	2007	G	T	G	A	T	T	C	T	G	A	T	T	T	T	T	T	T	T	2066
Db	1935	G	T	T	A	T	C	A	T	G	A	T	T	T	T	T	T	T	T	1994
Qy	2067	G	A	G	A	T	G	A	T	G	A	T	T	T	T	T	T	T	T	2126
Db	1995	G	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	2054
Qy	2127	A	A	C	C	C	G	T	T	G	A	C	T	G	A	T	T	T	T	2186
Db	2055	A	A	C	C	G	A	T	G	A	C	T	G	A	T	T	T	T	T	2114
Qy	2187	G	C	G	A	C	T	T	G	T	G	G	A	T	T	T	T	T	T	2246
Db	2115	G	C	T	G	A	C	T	T	G	T	G	G	A	T	T	T	T	T	2174
Qy	224																			

[illegible]

RESULT 2

```

PCT-US96-05800-1
; Sequence 1, Application PC/TUS9605800
; GENERAL INFORMATION:
; APPLICANT: OraVax, Inc.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05800
; FILING DATE: 23-APR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,041
; FILING DATE: 28-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA: 08/568,122
; APPLICATION NUMBER: 06-DEC-1995
; FILING DATE: 06-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/020001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2735 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US96-05800-1

```

Query Match 35.5%; Score 1021.4; DB 7; Length 2735;
 Best Local Similarity 66.5%; Pred. No. 6.6e-244;
 Matches 1634; Conservative 0; Mismatches 786; Indels 38; Gaps 9;
 QY 182 AAATTTTACACAGGAGTAAATAGGTGAACCTCACACCCAAAGACAGAAAAAGTTCTTGTG 241

Db	78	ACA	CTTTAAGAANTAGGAGAATGAGATGA	AACCTCA	CCCCCAAAAGAGTTAGATAAGTTGATG	137	
Qy	242	TTATAT	TATATGCGGCGAAGTGGCTAGAAAAG	CGCAAGCAGACAGAGGCTTAAAGCTCA	CAACAA	301	
Db	138	CTCCACT	AGCTGAGAAATTTGGCTTAAAAACG	CGCAAGAAAAAGGCATTAAGCTTAACTAT		197	
Qy	302	CCGGAAGCCAT	TGCTTACATTTAGTGGCCCATATTA	TATGGA	CGAAGCGGCCCTGTGGA	361	
Db	198	GTAGAAGCAGT	AGCTTTTGATTTAGTGCCCATATTA	TATGGA	GAAGCGAGAGCTGGTAAAAAG	257	
Qy	362	ACCGTTC	CCCCAGCTTATGGAGAGAGTGCAT	CTTTTGTGAAAAAAGATGAAGTAATG	CCCC	421	
Db	258	ACTCGCGT	GTAATTTGATGCAAGAGGCGGCAT	CTCTTTTAAAAACAGATGATGATGATG	AT	317	
Qy	422	GGGCTGGG	TATATGTTTCCCGATCTAGGTGTAGA	AGCCACCTTTCTCTGATGTCGAAAA		481	
Db	318	GGCGTGGC	AGCATGATCCATGAATGGGTATTA	TGAAGCGATGTTTCTGATGGGACTAAA		377	
Qy	482	CTTGTAAT	CTGGAATTTGGCCCATCGAAC	ACAGATGAGCACTTTCAAAGCGGCGGAAGTGA	AAAA	541	
Db	378	CTCGTAAC	CCGTGCATCCCTATTGAGGCGCAAT	TGGTAAATTTAGTTCTCTGGTGAGTTG	--	434	
Qy	542	TTTGGTT	CGCATAAAGACATCGAGCTCAATG	CGAGCGCAAGAGTAACCGA	ACTTTGAGTTT	601	
Db	435	TTCTTAAAAA	ATGAAGACATCACTATCA	CGAAGCGCAAAAACCGGTTAGCGTGA	AAAGTT	494	
Qy	602	ACTAATGA	AGGGCCTAAATCTCTTGCA	TGGGTAGCCATTTCCACTTCTTTGAAGCTA	AAAC	661	
Db	495	AAAAATGTT	GGCGACAGACCGGTTCAATCGG	CTCACACTTCTTTTGAAGTGAAT		554	
Qy	662	AAGGCAT	TAANAATTCGATCGTGAAAAAG	CCCTATGGCAAAAGCGCTAGATATTC	CCCTCTGGC	721	
Db	555	AGATGCT	GACTTTTGACAGAGAAAAA	ACTTTCGGTAAACGCTTAGACAT	TTCGAGCGGG	614	
Qy	722	AACAGCT	ACGATTTGGGGCAGGACAAAC	CCGCCAAAGTGCAGTTGATTCCTCTTGGTGGC		781	
Db	615	ACAGCGT	TAAGATTTTGAGCTGCGAAGAAAA	ATCCGTAGAAATTTGAATGACAT	TTCGCGGT	674	
Qy	782	AGTAAAAA	AGTGATGGCATGAACGGGCTTTG	TGAATAACATCCGCG--	-ATGAACGCCA-	837	
Db	675	AACAGAAGA	ATCTTTTGATTTAAGCATTTGGT	TGNTAGACACAGACAAACGAACGA		734	
Qy	838	-----TAAACA	-----TAAAGCGC-----	TTGA	CAAGGCG--AA--	ATCTCACG	872
Db	735	AAAATTC	CTTTTACACAGAGCTAAAGAGCG	TGGTGTTCATGCGCGCTAAAAGCG	GATGACAAC	794	
Qy	873	GATTT---	ATCAAGTAAGGAGACTCCCATGA	AAA---ATGAAAAACAAGAAATATG	TAAAT	926	
Db	795	TATGTAAAA	CAAAATTAAGGATGAAGAAATGA	AAAAAGATTTAGCAAAAAAGAAATATG	TTTCT	854	
Qy	927	ACCTACG	CGACCCCAAGCGGATAAGTGC	GTATAGGAGATACCGATCTTTTGGG	CAGAA	986	
Db	855	ATGTA	TGGTCTACTACAGCGGATAAAGT	GAGATTTGGCGGATACAGACTTGAT	TCGTGAA	914	
Qy	987	GTAGAAC	ATGACTATACCACTATGGCGA	AGAACTTTAAATTTGGCGCGGTAAA	AACTATC	1046	
Db	915	GTAGAAC	ATGACTACACCACTTATGCGA	AGAGCTTTAAATTCGGTGC	GGGTAAAAACCCCTA	974	
Qy	1047	CGTGAGG	GTATGGGTGAGCAATAGCCCTG	ATGAAAAACCCCTAGATTTAGTCA	CTACT	1106	
Db	975	AGAGAAG	CGCATGAGCCAACTCAAC	ACCCCTAGCAAAAGAGAGTTGAT	TTTAAATTAATCACT	1034	
Qy	1107	AACGGA	TGATTCGACTACCGGGATTTA	CAAAAGCCGACATTCGGGAT	TAAAAACGGC	1166	
Db	1035	AACGCTT	TAACTGTTGATTAACCGGTAT	TTATAAAGCGGATATTTGTTATTA	AGATGGC	1094	
Qy	1167	AAAATCC	ATGCGCATTTGGCAAGCGCAG	GAACCAAGGATCGCGCTTAAG	CCCTCAT	1226	
Db	1095	AAAATCC	TGCGCATTTGTAAGCGCGTAA	CAAGACATCAAGATGCGCT	TAAAAACAAT	1154	
Qy	1227	ATGGT	CTGGTGTGGGCACAGAAGCA	CTTAGCAGGGGAAGTTATGATTA	TTATCCGCTGGG	1286	
Db	1155	CTTAGCG	TAGGTCCTGCTACTGAAGCCT	TAGCCGGTGAAGGTTTGATCGT	TAACGGCTGGT	1214	

Qy	1287	GGAAATCGAATTCA	CACACACCCACTTCCTCTTTCTCCACAACAAATTCCTACCGCTCTAGCCAAT	1346
Db	1215	GGTATTGACACACACATCCACTTCATTTCACCCCAACAAATCCCTACAGCTTTTGCAAGC	1274	
Qy	1347	GGCGTTACACCAATGTTTGGAGGGCGACAGGTCTCTGTAGATGGCAGGATGGCACTACT	1406	
Db	1275	GGTGTAACAAACCAATGATTTGGTGTGGAAACCGGTCTCTGTATGGCACTAATTCGGCACTACT	1334	
Qy	1407	ATCACTCCGGGCAAAATGGAATCTTGACCGCAVGTGTGGCGCAGCAGAGAAGTAGTATCTTATG	1466	
Db	1335	ATCACTCCAGCGCAGAGAATTTTAAATGATGCTCAGAGCGGCTGAGAAATATTTCTATG	1394	
Qy	1467	AATGTGGCTTTTGGGCAAAAGGCAATPAGCTCTAGCAAAAAAACAATTTGTAGAACAACTA	1526	
Db	1395	AAUATTAGTTTCTTGGCTAAAGGTAAACGCTTCTAAACGATCCGAGCTTAGCCGATCAAATT	1454	
Qy	1527	GAAGCGGGCGGATTTGGTTTTTAATTTGCATGAAGACTGGGGCAACAACAAGTGGCATC	1586	
Db	1455	GAAGCGGGTCGGAATGGCTTTTAAATTTACAGAGACTGGGGCCCACTCTCTTCTGCAATC	1514	
Qy	1587	GATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTGTGTATCCACACCGATACA	1646	
Db	1515	AATCATCGCTTAGATGTTGCGGCAAAATACGATGTGCAAGTGCATATCCACACAGACACT	1574	
Qy	1647	GTCAAATGAGCGAGTATATGTAGATGACACACCTAAATGCAATGAACGGCGCGCATCCAT	1706	
Db	1575	TTGAATGAAGCGGTTGTGTAGAAGACACTATGGCTGCTATTGCTGAGCACTATGCGAC	1634	
Qy	1707	GCCTACACATTTAGGGAGGGGTGGAGGACACTCACCTGATGTATACCACTGGCAGGC	1766	
Db	1635	ACTTTCACACTGAAGGCGTGGCGGGGAGCAGCGCTCTGATATTAATAAGTAGCGCGT	1694	
Qy	1767	GAGCTCAATATCTACCCCTCTCCACACCCCACTATTCCTCTATACCTTAATACGGTT	1826	
Db	1695	GAACACAACATTTCTCCCGTTTCCACTAAACCCCAACATCCCTTTACCGTGAATACAGAA	1754	
Qy	1827	GCAGAAACATTTAGACATGCTCATGACATGCCACCACTAGACAAAAAGCATCCCGAGGAT	1886	
Db	1755	GCAGAGCACATGACATGCTTATGGTGTGCCACCACTTGGATAAAGCATTAAGAAGAT	1814	
Qy	1887	TTTCAATTTTCTCAAGCCGTATCCGCCCGGCTCTATTCGGCGTGAAGATGTCTCCAT	1946	
Db	1815	GTTTCAGTTTCGTGATTCAGAGATCCGCCCTCAAAACCATTCGGGTGAAGACACTTTGTCAT	1874	
Qy	1947	GATATGGTGTGATCGGATGACAAAGCTCCGATTCGCAAGCAATGGGCGTGCAGGCGAA	2006	
Db	1875	GACATGGGAATTTTCTCAATCAACAGTTTCGACTCTCAAGCGATGGCGCGTGTGGGTGAA	1934	
Qy	2007	GTGATTCCTCGAACTTGGCAGACTGCGGATAAGAAATAAAAAAGAAATTTGGTAGCTTCCT	2066	
Db	1935	GTATATCACTAGAACTTGGCAAAACAGCTGACAAAAAAGAAAGAAATTTGGCCGCTGAAA	1994	
Qy	2067	GAAGATGGCAAGATAACGATAATTTCCGCATTAAAGCGCTACACTTCCAAATACACTATC	2126	
Db	1995	GAAGAAAAAGCGATTAACGACAACTTCAGGATCAACCGCTACTTGTCTAAATACACCATT	2054	
Qy	2127	AACCCGCTTTGACCAACGGGCTGAGCGAGTATATCGGCTCTGTGGAAAGGGGCAAGATC	2186	
Db	2055	AACCCAGCGATCGCTCATATGGGAATTAGCGGATGTAGTTTCAGTAGAAGTGGGCAAGTG	2114	
Qy	2187	GCCGACTTGTGTGGTATCTGCTCTTTTGGCGTAAAAACCCAAAAATCGTGATCAAA	2246	
Db	2115	GCTGACTTGGTATTTGTGGATCGACGATTTTGGCGGTGAACCCAAACATGATCATCAAA	2174	
Qy	2247	GGCGGTATGGTGTCTCTCTGAAATGGGCGATTTCTAACCGCTGTGTGCCACTTCCCAA	2306	
Db	2175	GGCGGATTCATTCGGTTTAAGCCAAATGGCGATGCGGAACGCTTCTATCCCTACCCCAA	2234	
Qy	2307	CCGCTTTATTACCGCGAAATGTTTTGGGCGATCAACGCAAGGCGAAATTTGACACGAGATC	2366	
Db	2235	CCGCTTTATTACAGAGAAATGTTTCGCTCATCATGTTAAAGTAAATACGATGCAAAATC	2294	

Qy	2367	ACTTTTGTCTTCAAGTGCCTATGAAATGCGTGAAGAAAGAAAGCTGGCTTAGACGC	2426
Db	2295	ACTTTTGTCTCAAGCGGCTTATGACAAAGGCAATTAAGAAAGAAATTAGGACTTGAAGA	2354
Qy	2427	CAAGTCTTACCGGTCAAAAACCTGCCGTAAACATCACCAAGAAAGACATTCAGTTTCAACGAC	2486
Db	2355	CAAGTGTTCGCGGTAAAAAATTGCAAAATATCACTAAAAAAGACATGCANTTCAACGAC	2414
Qy	2487	AAAAACGGCAAAATCACCGTGCATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTC	2546
Db	2415	ACTACTGCTCACTTGAAGTCAATCTCGAACTTACCATGTGTTCGTGGATGGCAAGAA	2474
Qy	2547	TGCACCTCTAAACCCACTCGCAAGTGCTCTAGCCACGCGTACACTTCTCTCTAGG	2604
Db	2475	GTAACCTTCAAAACCAAGCAATAAAGTGAGCTTGGGCGCAACTCTTTTAGCATTTTCTAGG	2532
RESULT 3			
US-08-467-822-19			
; Sequence 19, Application US/08467822			
; Patent No. 5843460			
; GENERAL INFORMATION:			
; APPLICANT: Labigne, Agnes			
; APPLICANT: Sauerbaum, Sebastien			
; APPLICANT: Ferrero, Richard L.			
; APPLICANT: Thiberge, Jean-Michel			
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST			
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE			
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID			
; TITLE OF INVENTION: POLYPEPTIDES			
; NUMBER OF SEQUENCES: 44			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &			
; ADDRESSEE: Dunner			
; STREET: 1300 I Street, N.W.			
; CITY: Washington			
; STATE: D.C.			
; COUNTRY: USA			
; ZIP: 20005-3315			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patent In Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/467,822			
; FILING DATE: 06-JUN-1995			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/447,177			
; FILING DATE: 19-MAY-1995			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/432,697			
; FILING DATE: 02-MAY-1995			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Meyers, Kenneth J.			
; REGISTRATION NUMBER: 25,146			
; REFERENCE/POCKET NUMBER: 03495.0137-02000			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (202) 408-4000			
; TELEFAX: (202) 408-4400			
; INFORMATION FOR SEQ ID NO: 19:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 2619 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; MOLECULE TYPE: DNA (genomic)			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: 31..36			

OTHER INFORMATION: /standard_name= "Shine-Dalgarno"
OTHER INFORMATION: sequence.
FEATURE:
NAME/KEY: misc_feature
LOCATION: 756..759
OTHER INFORMATION: /standard_name= "Shine-Dalgarno"
OTHER INFORMATION: sequence.
US-08-467-822-19

Query Match 35.0%; Score 1008.8; DB 2; Length 2619;
Best Local Similarity 66.2%; Pred. No. 8.9e-241;
Matches 1634; Conservative 0; Mismatches 792; Indels 42; Gaps 8;

Qy	175	TGCTTTAAATTTAAACAAGGAGTATAGTGAACCTCACCCAAAGAGCAGAAAA	234
Db	13	TACCAATAGAAATTTCA-ATAAGGAGTTTAGGATGAACCTAACGCTTAAAGAACTAGACAA	71
Qy	235	GTCTCTGTATATATGCGCGGAAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCT	294
Db	72	GTTAATGCTCATTTATGCGGCAGATTGGCAGAGAACGCTTGGCGGTGTGTGAAACT	131
Qy	295	CAACCAACCCGAAGCCATTGCTTACATTAGTGCCTATATTATGACGAAAGCGCGCTGG	354
Db	132	CAATTACACGAAGCGGTCCGCTCATTTAGCGGCGTGTATGAAAGCGCGTGTGG	191
Qy	355	AAAAAAACCCTGCCAGCTTTATGGAAGAGTGATGCACCTTTTGTGAAAAAGATGAAGT	414
Db	192	TAATAAAAGCGTGGCGGATTTGATGCAAGAGGAGGAGCTTGGCTTAAATAAGAAAAATGT	251
Qy	415	AATGCCCGGGTGGTAAATATGGTTCCGATCTAGGTGTAGAACCCCTTTCCTGATGG	474
Db	252	GATGGACGGCGTAGCAAGCATGATTTCAATGAAGTGGGATTAAGCTAACTTCCCGGATGG	311
Qy	475	TACGAAATCTGTAATCTGTAATTTGCCCATCGAACCATGAGCACTTCAAAGCGGCGCA	534
Db	312	AACCAAGCTTGTAATCTCACTCCGGTAGAGTAATGGCAATTAGCCCCCGGCGCA	371
Qy	535	AGTGAAATTTGGTTGCGATAAAGACATCGAGCTCAATGCAAGGCAAGAAAGTAAACCGAACT	594
Db	372	GGT---CTTCTTAAAAATGAGGACATTACTATTAAACGCGCGCAAGAACCTTAGCTT	428
Qy	595	TGAGTTACTAAATGAAGGCTTAATCTCTGATGTTGGGTAGGATTCACACTCTTTTGA	654
Db	429	GAAAGTGAATAAAGCGCATCTGCTGTGTCAGGTGGGATCACTTTCCACTTTCTTGA	488
Qy	655	AGCTAACAGGCATAAATTTGATCGTGAAGAAAGCCATGCGCAAGCCCTAGATATTCC	714
Db	489	AGTGAATAAGCTTTGGACTTCGATTCGCGCAAAAGCTTTTGCAACGCTTAGACATTGC	548
Qy	715	CTTCGCAACACAGCTAGCGCATTTGGGCGAGGACAAACCCGCAAGGTGAGTTGATTCCTCT	774
Db	549	ATCTGGAACAGCGTGGCTTTGAAACCGGGAGGAAAAAGTGTGGNACTCATTTGACAT	608
Qy	775	TGTTGGCAGTAATAAAGTATGTCATGAACCGGCTTTGTGAATAACATCGCGAT---GA	831
Db	609	CGCGGGAATAAGCGCATCTATGCTTTAAATCTTTTGTGGATCGCAAGCCGATGCCGA	668
Qy	832	ACGCCATAAC-----ATAAGCGCTTGACAA-----GGCGAATCTCAGG	873
Db	669	TGTTAAAAAACTCGGCTTAAACCGCTTAAAGAAAAAGGTTTGGGTCTGTAAACTCGCG	728
Qy	874	ATTT---ATCAAGTAAGGAGACTC-----CCATGAAAAATGA---AAAAACAAGA	916
Db	729	TTGTGAAGCACTTAAGATAAACAATAAGGAAAAACCATGAAAAAGATTTTCAAGAAAGA	788
Qy	917	ATATGTAATAACCTTACGGACCCCAAGAGCGATTAAGTGCCTTAGGAGATACCGATCT	976
Db	789	ATATGTTTCTATGATGTTGCCACTACCGGGGATCGTGTAGACTCGCGGACACTGATTT	848
Qy	977	TTGGGCGAAGTAGACATCATATACCACTATGGCGAAGAACTTAAATTTGGCGCGG	1036
Db	849	GATCTTTAGAGTAGGATGATTTGACCACTTATGGTGAAGAGATCAAAATTTGGGGCGG	908

Qy	1037	TAAAACTATCCGTGAGGGTATGSGTCAGAGCAATAGCCCTGATGAAAAACACCCTAGATTT	1096
Db	909	TAAAACTATCCGTGATGGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTAGATTT	968
Qy	1097	AGTCATCACTAACCGGATGATTATCGACTACACCGGGATTTTACAAAGCCGACATTTGGAT	1156
Db	969	GGTCTCACTAACGCCCTCATTTGGACTATAGGGCAATTTACAAAGCCGACATTTGGAT	1028
Qy	1157	TAAAAACGCAAAATCCATGGCATTTGGCAAGGAGGAAAAACAAGACATGCAAGATGGCGT	1216
Db	1029	TAAAGACGCAAGATTTGCAGGCAATTTGGCAAGGAGGCAATTAAGACATGCAAGATGGCGT	1088
Qy	1217	AGCCCTCATATGTCGTGGGTGGGACAGAGCACTAGCAGGGGAGGATGATTTAT	1276
Db	1089	AGATAAATATCTTTGCTAGGTCCTGCTACAGAGGCTTTGGCAGCTGAGGGCTTGATTTG	1148
Qy	1277	TACCCCTGGGGAATTCGATTTACACACCCACTTCTCTTCTTCCACAACAATTCCTTACCGC	1336
Db	1149	AACCGCTGGTCATCGATACGCATATTCACITTTCTCTCTCCCAACAATCCCTACTGC	1208
Qy	1337	TCTAGCCAAATGGCGTTACAAACATGTTTGGAGCGGCAAGGTCCTGTAGATGGCAGAA	1396
Db	1209	TTTTGCCAGCGGGTTACAACCATGATTGGAGGAGGACAGGACCTGCGGATGGCAGAA	1268
Qy	1397	TGCGACTACTATCACTCCGGGCAATGNACTTGCACCGCATGTTGGCGGCGCAGCAAGA	1456
Db	1269	TGCGACCACTCACTCCCGGACGCGCTAATCTAAAGATGATGTTGCGTGACGCCGAAGA	1328
Qy	1457	GTATCTATGAATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTGT	1516
Db	1329	ATAACGCAATGATCTAGGCTTTTGGCTTAGGGGAATGTGCTTACGAACCTCTTTACG	1388
Qy	1517	AGAAACAAGTAGAAGCGGCGGATTTGGTTTAAATTTGCAATGAGACTGGGGCAACACC	1576
Db	1389	CGATCAGATTGAAGCAGGCGGATTTGGTTTAAATTTCCAGGAAGACTGGGGAGCACACC	1448
Qy	1577	AAGTCGATCGATCACTGCTTGAGCGTGGCAGATGAATAGATGTCAGATTTGTATCCA	1636
Db	1449	TGCACTATTCAACCACTGCTCAATGTCGCCGATGAATACGATGTGCAAGTGGCTATCCA	1508
Qy	1637	CACGATACACTCAATGAGGAGGATTTAGTATGACACCCCTAAATGCAATGAACGGCG	1696
Db	1509	CACGATACCTTAAACGAGGCGGCTGTGTAGAAGACACCTTAGAGGCGATTCGCCGGCG	1568
Qy	1697	CGCCATCCATGCTTACCACTTGAAGGAGCGGCTGAGGACACTCACTGATGTTATCAC	1756
Db	1569	CACATCTTACTTCCACACTGAAGGGCTGGGGTGACACGCTCCAGATGTTATCAA	1628
Qy	1757	CATGGCAGCGAGCTCAATATTTCTACCTCTCCACCCACCTATTTCCCTATACCAT	1816
Db	1629	AATGGCAGGGAAATTTAACTTCTACCGCTCTACTAACCCGACCATTTCTTTCACCAA	1688
Qy	1817	TAATACGTTTCAGNACACTTAGACATGCTCATGACATGCGCACCCCTAGACAAACGCAT	1876
Db	1689	AAACACTGAAGCCGAGCACATGGACATGTTAATGTTGTGTCACCACTTGGATAAAGAT	1748
Qy	1877	CGCGAGGATTTACAATTTTCTCAAGCGTATTCGCCCGGCTCTATCGCGCTCAAGA	1936
Db	1749	CAAGGAAGTGTGAGTTTGGATTCGAGGATTCGCCCCCAAACTATCGCGCTGAGAA	1808
Qy	1937	TGTGCTCCATGATATGGGTGTGATCGCATGACAAGCTCGGATTCGCAAGCAATGGGGCG	1996
Db	1809	CCAACTCCATGATGGGATCTTTCTTCTATCACAGCTCCGACTCTCAGGCTATGGGACG	1868
Qy	1997	TGCGAGGAAATGATCTCTCGAACTTGGCAGACTCGGGATTAAGATATAAAGAAATTTGG	2056
Db	1869	CGTAGGCGAGGTGATCAACGCACTTTGGCAGACAGCAGACAAAAAAGAGTTTGG	1928
Qy	2057	TAAAGTCTCTCAAGTAGCAAGATAACGATAATTTCCGCAATTTAGCGCTACATCTCCAA	2116
Db	1929	CGCTTGAAAGGAGGAAAAAGGCGATAACGACACTTCCGCATCAACACGCTACATCTCAA	1988
Qy	2117	ATACACTATCAACCCCGCTTTTGACCCACCGGCTGAGCGAGTATATCGGCTCTGTGGAAGA	2176

```
Db 1989 ATACCATCAACCCCGGATCGGCATGGATTTCTGACTATGTGGCTCTGTGGAAGT 2048
Qy 2177 GGGCAAGATCCCGCATTTGGTGTGGATTCCTGCGCTTTTGGCGTAAACCCAAAT 2236
Db 2049 GGGCAATACCGACCTCGTGTGGATTCGGCTTTCTTGGCATTAAGCCCAATAT 2108
Qy 2237 CGTGATCAAGCGGTATGGTGTCTTCTGAAATGGCGATCTAACGGCTGTGTC 2296
Db 2109 GATTAATTAAGCGGATTTATTGGGCTCTCTCAAAATGGCGATGCCAATGGCTTATTC 2168
Qy 2297 CACTCCCAACCGGTTATTATCCGGAATCTTTGGGCATCAGCGAAGCGGAATTTGA 2356
Db 2169 CACCCCTCAGCCGCTTATTACCGTGAATTTTGGACACCATGGGAAAAACAATTCGA 2228
Qy 2357 CACGAGCATCACTTTTGTTCCTCAAGTCGCTATGAAATGGCGTGAAGAAAAGCTGGG 2416
Db 2229 CACCAATATCACTTCGTGTCCTCAAGCGGCTTACAGCGCAGGATCAAGAAAGACTAGG 2288
Qy 2417 CTTAGAGCGCAAGTTCTACCGGTCAAAAATCGCGTAACTACCAAGAAAGACTTCAA 2476
Db 2289 GCTAGATCGCGCGCAGCCAGTGAATACTGTGCAATATCACTAAAGAGGACCTCAA 2348
Qy 2477 GTTCAAGCAAAACGCAAAATCACCGTCGATCCGGAACCTTCGAGGCTTTGTAGA 2536
Db 2349 ATTCAACGATGTGACCGCACATATTGATGTCAACCTTGAAACCTATAAGGTGAAGTGA 2408
Qy 2537 TGGCAAACTCTGCACCTCTAAACCCACCTCCCAAGTGCCTCTAGCCACGCGCTACACTTT 2596
Db 2409 TGGCAAGAGTAACCTCTAAAGCAGCAGATGAATTGAGCCTAGCGCAACTTTATATTT 2468
Qy 2597 CTTCTAGG 2604
Db 2469 GTTCTAGG 2476
```

RESULT 4

```
US-08-432-697-19
; Sequence 19, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
```

```
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; US-08-432-697-19
Query Match 35.0%; Score 1008.8; DB 3; Length 2619;
Best Local Similarity 66.2%; Pred. No. 8.9e-241;
Matches 1634; Conservative 0; Mismatches 792; Indels 42; Gaps 8;
Qy 175 TGCCTTTAAATTTAAACAAGAGGATTAAGTGAACCTCACACCCAAAGAGCAAGAAA 234
Db 13 TACCAATAGAAATTTCA-ATAAGGAGTTTAGGATGAACCTTAACGCTTAAAGAACTAGACAA 71
Qy 235 GTTCTTGTATTATATGCGGCGAAGTGGCTAGAAAGCCAAAGCAGAGGGCTTAAAGCT 294
Db 72 GTTAATGCTCCATATATGCGGCGAGATTGGCAGAAAGAACCTTGGCGCGTGGTGTGAACCT 131
Qy 295 CAACCAACCCGAAAGCCATTGCTTTACATTAGTGTCCCATATTATGGACGAAAGCGCGCTGG 354
Db 132 CAATTACACCGAAGCGGTCCGCTCATTTAGCGGCGTGTGTGTAAGAAAGCGCGTGTG 191
Qy 355 AAAAAAACCGTGTGCCAGCTTATGGAAGAGTGCATGCACTTTTGGAAAAAGATGAAGT 414
Db 192 TAATAAAAGCGTGGCGGATTTGATCAAGAGCAGGAGCTTGGCTTTAAAAAAGAAAAATGT 251
Qy 415 AATGCCCGGGTGGTAAATATGTTTCCCGATCTAGTGTAGAGCCACCTTTCCTGATGG 474
Db 252 GATGGACCGGTAGCAAGCATGATTCATGAAGTGGGGATTGAAGCTAACTTCCCGCATGG 311
Qy 475 TACGAAACTTGTAACTGTGAATTGGCCCATTCGAAACAGATGAGCACTTCAAAGCGGCGCA 534
Db 312 AACCAAGCTTGTAACTATCCACACTCCGCTAGAGGATATGCAAAATTAGCCCCCGCGCA 371
Qy 535 AGTGAATTTGGTTCGGATAAAGACATCGAGTCAATGAGGCAAGAAAGAGTAACCGAACT 594
Db 372 GGT---CTTCTTAAAAAATGAGGACATTACTATTAAACGCGGCAAGAAAGCACTTAGCTT 428
Qy 595 TGAGTTACTAATGAAGGCGCTAAATCTTGCATGTGGGTAGCCATTTCCACTTCTTTGA 654
Db 429 GAAAGTGAATAAAGGCGATCGCTGTGAGGTGGGATCACAATTTCCACTTCTTCGA 488
Qy 655 AGCTAAACAAGGCACCTAAAAATTCGATCGTGAAGAGCTTATGGCAAGCGCTAGATATTC 714
Db 489 AGTGAATAGCTCTTGGACTTCGATCGCGCAAAAAGCTTTTGCAAACGCTAGACATTC 548
Qy 715 CTCTGGCAACACGCTACGCTAGCATTTGGGGCAGGACAAACCCGCAAGTGCAGTTGATTCCT 774
Db 549 ATCTGNAACAGCGGTGCGCTTTTGAACCCGGGAGGAAAAAGTGTGGAACCTCATTGACAT 608
Qy 775 TGGTGGCAGTAAAAAGTGAATGGCATGAACGGGCTTGTGTAATAACATCGCGAT---GA 831
Db 609 CGCGGGAATAAGCGCATCTATGGCTTTTAACTTTTGTGGTGGATCGCCAAAGCCGATGCCGA 668
Qy 832 ACGCCATAAAC-----ATAAGCGCTTGAACA-----GGCGAAATCTCACGG 873
Db 669 TGGTAAAAAATCGCGCTTAAACCGCGCTAAGAAAGAGTTTGGGTCTGTAACTGCGG 728
Qy 874 ATTT--ATCAAGTAAGGAGACTC-----CCATGAAATGA---AAAAACAAGA 916
```


Db 729 TTGTGAAGCGACTAAGATAAAGTAAGAAACCAATGAAGAAAGATTTCACGAAAGA 788
Qy 917 ATATGTAATAACCTACCGGACCCACCAAGGCGATAAAGTGGCGCTTAGGAGATACCGATCT 976
Db 789 ATATGTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 848
Qy 977 TTGGGCGAGAGTAGAACAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1036
Db 849 GATCTTGAAGTGGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 908
Qy 1037 TAAACTATCGTGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1096
Db 909 TAAACTATCGTGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 968
Qy 1097 AGTCATCACTAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1156
Db 969 GTGTCTCACTAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1028
Qy 1157 TAAAGCGGCAAAATCCATGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1216
Db 1029 TAAAGCGGCAAAATCCATGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
Qy 1217 AAGCCCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1276
Db 1089 AGATAAATCTTGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1148
Qy 1277 TACCGCTGGGGGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1336
Db 1149 AACCGCTGGGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1208
Qy 1337 TCTAGCAATGGCGTTACAAACCATGATGATGATGATGATGATGATGATGATGATGATGAT 1396
Db 1209 TTTTGGCAGCGGTTACAAACCATGATGATGATGATGATGATGATGATGATGATGATGAT 1268
Qy 1397 TCGCATCTATCTACCTCGGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1456
Db 1269 TCGCACCACCATCTCCCGGACCGCTTAATCTTAAAGATGATGATGATGATGATGATGAT 1328
Qy 1457 GTATTTCTAAGTGGGCTTTTGGGCAAGGCAATGATGATGATGATGATGATGATGATGAT 1516
Db 1329 ATAGCCATGATCTAGGCTTTTGGCTTAAAGGCAATGATGATGATGATGATGATGATGAT 1388
Qy 1517 AGAACAGTAGAAGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1576
Db 1389 CGATCAGATTGAAGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1448
Qy 1577 AAGTGGATCGATCACTGCTTGAAGCGGCGGATGATGATGATGATGATGATGATGATGAT 1636
Db 1449 TGCAGCTTATTCACCATCTGCTCAATGTCGCGATGATGATGATGATGATGATGATGATGAT 1508
Qy 1637 CACCGATACAGTCAATGAGCGGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1696
Db 1509 CACCGATACCTTAAACAGCGGCGGCTGTGATGATGATGATGATGATGATGATGATGATGAT 1568
Qy 1697 CGCATCTCATGCTTACCATGATGAGCGGCGGATGATGATGATGATGATGATGATGATGAT 1756
Db 1569 CACCATCTACCTTCACTAGTGAAGCGGCGGATGATGATGATGATGATGATGATGATGATGAT 1628
Qy 1757 CATGGCAGCGGAGCTCAATATCTACCTCTCCACCAACCGGCGGATGATGATGATGATGAT 1816
Db 1629 AATGGCAGCGGGAATTTAACTTCTACCGGCTTACTAAACCGGCGGATGATGATGATGATGAT 1688
Qy 1817 TAATAGGTTGCAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1876
Db 1689 AAACACTGAAGCGGAGCACAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1748
Qy 1877 CCGCAGGATTTTCAATTTTCTCAAGCGGCTATCCGCGGCGGCTTATCCGCGGCTGAAGA 1936
Db 1749 CAAGGAAGATGTCAGTTTCCGATTCGAGATTCGCGGCGGCTTATCCGCGGCTGAAGA 1808
Qy 1937 TGTGCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1996

Db 1809 CCAACTCCATGACATGGGATCTTTTCTATCACCAGCTCCGACTCTCAGGCTATGGGACG 1868
Qy 1997 TCGAGCGAAGTGAATTCCTCGAACTTTGGCAGACTCGGATAGAAATAAATAAAGATTTGG 2056
Db 1869 CGTAGCGAGGTGATCACACGCACTTTGGCAGACAGCAGACAAAAACAAAAAGATTTGG 1928
Qy 2057 TAAGCTTCTGAAGTGGCAAGATAACGATATTTCCGATTTAAGCGCTACATCTCTCAA 2116
Db 1929 GCGCTTGAAGAGGAAAAAGGCGATAAACCACTTCCGATCAAAAGCTACATCTCTAA 1988
Qy 2117 ATACACTATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGA 2176
Db 1989 ATACACTATCAACCCCGGATCGGATGAGGATTTCTGACTATGTGGCTCTGTGGAAGT 2048
Qy 2177 GCGCAAGATCGCGACTTTGGTGTGGAATCTGCTTTTGGCGTAAAAACCCAAAT 2236
Db 2049 GGGCAATACCGCGACTCTGCTTTGGAGTCCGCTTTCTTTGGCATTAAGCCCAATAT 2108
Qy 2237 CGTATCAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2296
Db 2109 GATTATTAAGGCGGATTTAATGCGCTCTCTCAATGGGCGATGCCAATGCTCTATTCC 2168
Qy 2297 CACTCCCAACCGGCTTTATCCGCGAATGTTTGGGCAATCGGCAAGCGGCAATTTGA 2356
Db 2169 CACCCCTCAGCCGCTATTACCGGAAATGTTTGGGCAATCGGCAAGCGGCAATTTGA 2228
Qy 2357 CACCAAGCATCACTTTTGGTTCAGGCTATGAAATGGCGTGAAGAAAGCTGGG 2416
Db 2229 CACCAATATCACTTTGCTGCCAGCGCTTACAAGGCGGATCAAAGAGACTAGG 2288
Qy 2417 CTTAGAGCGCAAGTTTACCGGTCAAAAACCTGCGGTAAACATCAACAGAAAGACTCAA 2476
Db 2289 GCTAGTCCGCGGCGACCGCACTGCGCAATATCACTAAAAAGGACCTCAA 2348
Qy 2477 GTTCAACGCAAAACCGCAAAATCACCGTCAATCCGAAACCTTCGAGGCTTTGTAGA 2536
Db 2349 ATTCAACGATGACCGCACAATTTGATGATCAACCTCGAAACCTATAAGGTGAAGTGA 2408
Qy 2537 TGGCAAACTCTGACCTCTAAACCCACCTCGCAAGTGCCTTAGCCGCGCTACACTTT 2596
Db 2409 TGGCAAGAGGTAACTCTTAAGCAGCAGATGAATGAGCGCTAGCGCACTTTATATTT 2468
Qy 2597 CTTCTAGG 2604
Db 2469 GTTCTAGG 2476

RESULT 5

US-08-466-248-19
; Sequence 19, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno"
; OTHER INFORMATION: sequence."
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno"
; OTHER INFORMATION: sequence."
; US-08-466-248-19
;
; Query Match 35.0%; Score 1008.8; DB 3; Length 2619;
; Best Local Similarity 66.2%; Pred. No. 8.9e-241;
; Matches 1634; Conservative 0; Mismatches 792; Indels 42; Gaps 8;
;
; Qy 175 TGCCTTTAAATTTAAACAAGGAGTAAATAGGTGAAACTCACACCAAGAGCAAGAAAA 234
; Db 13 TACCAATAGAATTCATATAGGAGTTAGGATGAACTAAGCCCTTAAGAACTAGACAA 71
;
; Qy 235 GTTCTTGTTATATTATGCGGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCT 294
; Db 72 GTTAATGCTCCATTATGCGGGCAGATTGGCAGAGAAACGCTTGGCGCGTGTGTAACCT 131
;
; Qy 295 CAACCAACCCGAGCATGCTTATAGTCCCATATTATGAGCAAGCGCGCGGTGG 354
; Db 132 CAATATACACCAAGCGGTGCGGCTCATTAGCGGGCGTGTGATGTAAGAAAGCGCGTGTGG 191
;
; Qy 355 AAAAAAACCGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGAAGAAAGATGAAGT 414
; Db 192 TAATAAAGCGTGGCGGATTTGATGCAAGAGCGAGGACTTGGCTTAAAAAGAAATGT 251
;
; Qy 415 AATGCCCGGGTGGTAAATATGGTTCCCGATCTAGGTGTAGAGCACTTTTCCTGATGG 474
; Db 252 GATGACGCGGTAGCAAGCATGATTATGAAAGTGGGATTAAGCTAACTTCCCGCATGG 311
;
; Qy 475 TACGAACTTGTAACTGTGAATGGCCCATCGAACAGATGAGCACTTCAAGCGGGCGA 534
; Db 312 AACCAAGCTTGTAACTATCCACACTCCGGTAGAGGATTAATGCAAAATAGCCCGCGGA 371
;
; Qy 535 AGTGAATTTGGTTCGATTAAGACATCGAGCTCAATGCAAGAGAGTAACCGAACT 594
; Db 372 GGT---CTTCTTAAAAATGAGACATTAATAATACGCGCGGCAAGAGCCATTAGCTT 428
;
; Qy 595 TGAGTTACTAATGAAGGCGCTTAAATCTTGTGATGTGGGTAGCCATTTCCTTTTGA 654
; Db 175 TGCCTTTAAATTTAAACAAGGAGTAAATAGGTGAAACTCACACCAAGAGCAAGAAAA 234
; Db 13 TACCAATAGAATTCATATAGGAGTTAGGATGAACTAAGCCCTTAAGAACTAGACAA 71
;
; Qy 235 GTTCTTGTTATATTATGCGGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCT 294
; Db 72 GTTAATGCTCCATTATGCGGGCAGATTGGCAGAGAAACGCTTGGCGCGTGTGTAACCT 131
;
; Qy 295 CAACCAACCCGAGCATGCTTATAGTCCCATATTATGAGCAAGCGCGCGGTGG 354
; Db 132 CAATATACACCAAGCGGTGCGGCTCATTAGCGGGCGTGTGATGTAAGAAAGCGCGTGTGG 191
;
; Qy 355 AAAAAAACCGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGAAGAAAGATGAAGT 414
; Db 192 TAATAAAGCGTGGCGGATTTGATGCAAGAGCGAGGACTTGGCTTAAAAAGAAATGT 251
;
; Qy 415 AATGCCCGGGTGGTAAATATGGTTCCCGATCTAGGTGTAGAGCACTTTTCCTGATGG 474
; Db 252 GATGACGCGGTAGCAAGCATGATTATGAAAGTGGGATTAAGCTAACTTCCCGCATGG 311
;
; Qy 475 TACGAACTTGTAACTGTGAATGGCCCATCGAACAGATGAGCACTTCAAGCGGGCGA 534
; Db 312 AACCAAGCTTGTAACTATCCACACTCCGGTAGAGGATTAATGCAAAATAGCCCGCGGA 371
;
; Qy 535 AGTGAATTTGGTTCGATTAAGACATCGAGCTCAATGCAAGAGAGTAACCGAACT 594
; Db 372 GGT---CTTCTTAAAAATGAGACATTAATAATACGCGCGGCAAGAGCCATTAGCTT 428
;
; Qy 595 TGAGTTACTAATGAAGGCGCTTAAATCTTGTGATGTGGGTAGCCATTTCCTTTTGA 654
; Db 175 TGCCTTTAAATTTAAACAAGGAGTAAATAGGTGAAACTCACACCAAGAGCAAGAAAA 234
; Db 13 TACCAATAGAATTCATATAGGAGTTAGGATGAACTAAGCCCTTAAGAACTAGACAA 71
;
; Db 429 GAAAGTGAATAAAGGGGATCGTCTCTGCGAGGTGGGATCACATTTCCACTTCTTGA 488
; Qy 655 AGCTAAACGAGCCTATAAATTCGATCGTGAAGAGCCTATGGCAACCGCTAGATATTC 714
; Db 489 AGTGAATAAGCTCTTGGAGCTTCGATCGCGCAAAAAGCTTTTGCAACCGCTAGACATTC 548
;
; Qy 715 CTCTGGCAACGCTTACGATCGGATGGGGGAGGACAAACCCCGCAAGTGCAGTTGATCTCT 774
; Db 549 ATCTGGAACAGCGGTGCGCTTTGAAACCCCGGGAGGAAAAAAGTGTGGAACTCATTTGACAT 608
;
; Qy 775 TGGTGGCAGTAAAAAAGTATGGGATGAACGGCTTGTGAATAACATCCGGAT---GA 831
; Db 609 CGCGGGGATTAAGCGCATCTATGGCTTTAATTTCTTTGGTGGATCGCCAAGCCGATGCCGA 668
;
; Qy 832 AGCCATAAAC-----ATAAGCGCTTTGACAA-----GGCGAAATCTCACGG 873
; Db 669 TGGTAAAAAACTCGGCTTAAAAACGCGCTAAAGAAAAAGGTTTGGGCTCTGTAACCTGCGG 728
;
; Qy 874 ATTT--ATCAAGTAAGGAGACTC-----CCATGAAAAATGA---AAAAACAAGA 916
; Db 729 TTGTGAAGCGACTAAAGATTAACAATAAGGAAAAAACCATGAAAAAGATTTTCAGAAAAA 788
;
; Qy 917 ATATGTAATACCTACGGACCCACCAAGCGGATTAAGTGCCTTAGGAGATACCGATCT 976
; Db 789 ATATGTTTCTATGATGTGCTCCCACTACCGGGGATCGTTAGACTCGGCGACTGATTT 848
;
; Qy 977 TTGGGCGAGAAGTAGAACATGACTATATACCACTATGGCGAAGAACTTTAAATTTGGCGCGG 1036
; Db 849 GATCTTAGAAGTGGAGCATGATTGACCACTTATGCTGAAGAGATCAATTTGGGGGCGG 908
;
; Qy 1037 TAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACCCCTAGATTT 1096
; Db 909 TAAAACTATCCGTGATGGGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTAGATTT 968
;
; Qy 1097 AGTCATCACTAACGCGATGATTTATCGACTACACCGGATTTTACAGCCGACATCCGGAT 1156
; Db 969 GGTGCTCACTAACGCGCTCATTTGGAGCTATACGGGCAATTTACAAAGCCGACATGGGAT 1028
;
; Qy 1157 TAAAAACGCGAAAAATCCATGGCATTTGGCAAGGAGGAAAAACAAGGACATGCAAGATGGCGT 1216
; Db 1029 TAAAGCGCGAGATTTGAGGATTTGGCAAGGAGGCAATAGGACATGCAAGATGGCGT 1088
;
; Qy 1217 AAGCCCTCATATPGTGTGGGTGTGGGCAAGAGCACTAGCAGGGGAAAGGTATGATAT 1276
; Db 1089 AGATAATAATCTTTCGTAGGTCTGCTACAGAGGCTTTGGCAGCTGAGGGCTGATTTG 1148
;
; Qy 1277 TACCGCTGGGGAATCGATTACACACCACTCTCTTCTCCCAACAATTCCTTACCGC 1336
; Db 1149 AACCCGCTGGTGGCATCGATACGATATTCATTTATCTCTCCCAACAATTCCTTACTGC 1208
;
; Qy 1337 TCTAGCCATGGCGTTTACAAACCATGTTTGGAGGCGCACAGGTCCTGTAGATGGCAGAA 1396
; Db 1209 TTTTCCAGCGGGTTTACAAACCATGATTGAGAGAGGCAAGGACCTGCGGATGGCAGAA 1268
;
; Qy 1397 TCGGACTACTTACCTCCCGGCAAAATGGAACCTTGCACCGCATGTTGCGCGCAGCAGAGA 1456
; Db 1269 TCGGACCACTACCTCCCGGACGCGCTAATCTAAAGAGTATGTTGCGTGCAGCCGAAGA 1328
;
; Qy 1457 GTATCTTATGAATGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTTGT 1516
; Db 1329 ATACGCCATGAATCTAGGCTTTTGGCTAAGGGGAAATGTTCTTACGAACCCCTCTTACG 1388
;
; Qy 1517 AGAACAAAGTAGAAGCGGGCGGATTTGTTTAAATTTGATGAAGATGCGGGCACACACC 1576
; Db 1389 CGATCAGATTGAAGCAGGGGCGGATTTGTTTAAATTTCCAGAAAGCTGGGGAAGCACACC 1448
;
; Qy 1577 AAGTGCATTCATCTGCTTGAGCGTGGCAGATGAATAGATGTGCAAGTTGTATTCGA 1636
; Db 1449 TCGAGCTATTACCACTGCTCAATGTCGCGCATGAATACGATGTCAAGTGGCTATCCA 1508
;
; Qy 1637 CACCGATACAGTCAATGAGGCGAGTTATGTAGATGACACCTTAATGCAATGAACGGGCG 1696
; Db 1509 CACCGATACCTTAAACGAGGCGGCTGTGTAGAGGACACCCCTAGAGGCGGATTTGCGGCGG 1568
```

```

Qy 1697 CGCCATCATGCTTACCATTTAGGAGCGGTGGAGGACCTCCTGATGTTATCAC 1756
Db 1569 CACCATCATACCTTCCACACTGAAGGGGCTGGGGGTGGACACGCTCCAGATGTTATCAA 1628
Qy 1757 CATGGCAGGGAGCTCAATATTTCTACCTCTCTCCACACCCCTCATTTCCCTATACCAT 1816
Db 1629 AATGGCAGGGAGTTTAACTTCTACCTCTCTCTCCACACCCCTCATTTCCCTATACCAT 1688
Qy 1817 TAATACGGTTTGCAGAACATTTAGACATGCTCATGACATGCCACACCTAGAACACGCAT 1876
Db 1689 AAACACTGAAGCCGAGCAGACATGGACATGTTAATGGTGTGCCACCTTTGGATATAAGTAT 1748
Qy 1877 CCGGAGGATTTACATTTCTTCAAGCCGTATCGCCCCGGCTCTATCGCGGCTGAAGA 1936
Db 1749 CAAGGAAGATGTGCAGTTTGCAGTTTGCAGATTTCGCCCTTCCGCCCAAACTATCGCGGCTGAAGA 1808
Qy 1937 TGTGCTCATGATATGCTGTGATCGGATGACAGCTCGGATTCGCAAGCAATGGGGCG 1996
Db 1809 CCAACTCCATGACATGGGATCTTTCTTATCACCAGCTCCGACTCTCAGGCTATGGGACG 1868
Qy 1997 TGCAGGCGAAGTGAATCTCTCGAACTTGGCAGACTCGGATAGAAATAAAGAAATTTGG 2056
Db 1869 CGTAGGCGAGTGATCACACGACTTGGCAGACGACAGACAAACAAAGAGTTGG 1928
Qy 2057 TAAGCTTCTGAAGATGGCAAGATAACGATPAATTTTCGCATTAAAGCGCTACATCTCAA 2116
Db 1929 CGCTTGAAGAGGAAAGAGCGATAACGACAACTTCGCGATCAAGCGCTACATCTCTAA 1988
Qy 2117 ATACACTATCAACCCCGCTTACCCAGCGGTGAGCGATATATCGCTCTGTGGAGA 2176
Db 1989 ATACACTATCAACCCCGGTATCGCGATGCGGATTTCTGACTATGTGGCTCTGTGGAGT 2048
Qy 2177 GGGCAAGATCGCGACTTGTGTGTGCTGCTCTCTGAAATGGCGGATTTCTAAGCGCTGTGGCC 2236
Db 2049 GGGCAATATCGCGACTTGTGTGTGCTGCTCTCTGAAATGGCGGATTTCTAAGCGCTGTGGCC 2108
Qy 2237 CGTGATCAAGCGGATATGTTGCTCTCTGAAATGGCGGATTTCTAAGCGCTGTGGCC 2296
Db 2109 GATTATTAAGCGGATTTATTTGGCTCTCTCAATGGCGGATGCAATGCTGTATTC 2168
Qy 2297 CACTCCCAACCGGTTTATTAACCGGAAATGTTTGGGATCAAGCGGAAATTTGA 2356
Db 2169 CACCCCTCAGCGCTTATTAACCGTGAATGTTTGGACACCATGGGAAACAAATTCGA 2228
Qy 2357 CACGAGATCACTTTGTTTCCAAAGTCGCTATGAATGCGGTGGAAGAAAGCTGGG 2416
Db 2229 CACCAATATCACTTTGTTTCCAAAGTCGCTTCAAGCGGATTCAGGCAAGAACTAGG 2288
Qy 2417 CTTAGAGCGCAAGTTCTACCGGTCAAAATCTGCGGTAAACATCAACCAAGAAAGCTTCAA 2476
Db 2289 GCTAGATCGGCGGACCGGATGAAATCTGCGCAATATCACTAAGGACCTCAA 2348
Qy 2477 GTTCAACGACAAACGGCAAAATACCGGTGATCGGAAACCTTCGAGGTCTTTGAGA 2536
Db 2349 ATTCAAGATGTGACGACATATTTGATGTCACCTTGAACCTTAAAGTGAAGTGA 2408
Qy 2537 TGCAAACTCTGACCTCTTAAACCCACTCGCAAGTGCCTCTAGCCAGCGCTACATTT 2596
Db 2409 TGGCAAGAGGTAACCTCTAAAGCAGCAGATGATTTAGCGCTAGCGCACTTTATAATTT 2468
Qy 2597 CTTCTAGG 2604
Db 2469 GTTCTAGG 2476

```

RESULT 6

US-09-431-705-1
; Sequence 1, Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Londono-Arcila, Patricia

```

; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; FILE REFERENCE: 06132/06001
; CURRENT APPLICATION NUMBER: US/09/431,705
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4824
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: includes sequences from Helicobacter pylori,
; OTHER INFORMATION: Salmonella typhimurium, and Escherichia coli
; NAME/KEY: CDS
; LOCATION: (2)...(31)
; NAME/KEY: CDS
; LOCATION: (41)...(61)
; NAME/KEY: CDS
; LOCATION: (65)...(799)
; NAME/KEY: CDS
; LOCATION: (803)...(2512)
; NAME/KEY: CDS
; LOCATION: (2516)...(2692)
; NAME/KEY: CDS
; LOCATION: (2696)...(2896)
; NAME/KEY: CDS
; LOCATION: (2900)...(3322)
; NAME/KEY: CDS
; LOCATION: (3326)...(3385)
; NAME/KEY: CDS
; LOCATION: (3389)...(3406)
; NAME/KEY: CDS
; LOCATION: (3410)...(3466)
; NAME/KEY: CDS
; LOCATION: (3470)...(3598)
; NAME/KEY: CDS
; LOCATION: (3602)...(3661)
; NAME/KEY: CDS
; LOCATION: (3665)...(3697)
; NAME/KEY: CDS
; LOCATION: (3701)...(3769)
; NAME/KEY: CDS
; LOCATION: (3773)...(3817)
; NAME/KEY: CDS
; LOCATION: (3821)...(3844)
; NAME/KEY: CDS
; LOCATION: (3848)...(3889)
;
US-09-431-705-1

```

```

Query Match      34.9%; Score 1005.8; DB 3; Length 4824;
Best Local Similarity 65.8%; Pred. No. 6.6e-240;
Matches 1640; Conservative 0; Mismatches 812; Indels 39; Gaps 10;

Qy 150 AACGAAATTTAGGACTATATCCATTTGCTTTTAAATTTAAACACAAG-GAGTAATAGGTG 208
Db 29 AAATGAATCTGACGTACACAGCAATTTAGATATTAATCATCCACAGAGAGATCTCCATG 88
Qy 209 AAATCTCACCAACCAAGAGCAAGAAAGTTCTTTGTTATATTTATGCGGCGAAGTGGCTAGA 268
Db 89 AAATCTCACCAACCAAGAGTTAGATAAGTTGATGTCTCACTACGCTGGAGAAATGGCTAAA 148
Qy 269 AAGCGCAAGCAGAGCGGCTTAAAGCTCAACCAACCGGAGCCATTTCTTACATTAGTGCC 328
Db 149 AAACGCAAGAAAGAGGCAATTAAGCTTAATGTAGTAAGCAGTAGCTTTGATTAGTGCC 208
Qy 329 CATATTATGGACGAGCGCGCTGGAAGAAACCGTTTGGCCAGCTTATGGAAGAGTGC 388
Db 209 CATATTATGGAGAGCGGAGAGCTGTTAAAGACTGCGCTGAATTTGATGCAAGAGGG 268
Qy 389 ATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGTGGGTAATATGTTCCCGATCTA 448

```



```
QY 2574 CCTCTAGCCAGCGCTACACATCTTCTCTAGG 2604
| | | | | | | | | | | | | | | | | |
Db 2486 AGCTGGCGCACTCTTTAGCAATTTCTAGG 2516
| | | | | | | | | | | | | | | | | |

RESULT 7
US-09-431-705-19
; Sequence 19 Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Londono-Arcila, Patricia
; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; vaccination against Helicobacter infection
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431.705
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 4824
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: includes sequences from Helicobacter pylori,
; Salmonella typhimurium, and Escherichia coli
; NAME/KEY: CDS
; LOCATION: (3893)...(3934)
; NAME/KEY: CDS
; LOCATION: (3938)...(4027)
; NAME/KEY: CDS
; LOCATION: (4031)...(4285)
; NAME/KEY: CDS
; LOCATION: (4289)...(4300)
; NAME/KEY: CDS
; LOCATION: (4304)...(4408)
; NAME/KEY: CDS
; LOCATION: (4412)...(4471)
; NAME/KEY: CDS
; LOCATION: (4475)...(4598)
; NAME/KEY: CDS
; LOCATION: (4592)...(4669)
; NAME/KEY: CDS
; LOCATION: (4673)...(4711)
; NAME/KEY: CDS
; LOCATION: (4715)...(4774)
; NAME/KEY: CDS
; LOCATION: (4784)...(4824)
; US-09-431-705-19

Query Match 34.98; Score 1005.8; DB 3; Length 4824;
Best Local Similarity 65.88; Pred. No. 6.6e-240;
Matches 1640; Conservative 0; Mismatches 812; Indels 39; Gaps 10;

QY 150 AACGAATATAGGACTAATCCCATTCCTTTAAATTTAAACACAAG-GAGTAATAGGTG 208
| | | | | | | | | | | | | | | | | |
Db 29 AATGATCTGAGCTACACAGCAATTTAGATATATATCATCCACAGAGATCTCCATG 88
| | | | | | | | | | | | | | | | | |

QY 209 AAATCACCACCCAAAGAGCAAGAAAAGTTCTTGTATATATTATGCGGCGAAGTGGCTAGA 268
| | | | | | | | | | | | | | | | | |
Db 89 AAATCACCACCCAAAGAGTTAGATAGTTGATGCTCCACTACGCTGGAGAATTGGCTAAA 148
| | | | | | | | | | | | | | | | | |

QY 269 AACGGCAAGCAGGCTTAAAGCTCAACCAACCCGAGCATTGCTTACATTAGTGCC 328
| | | | | | | | | | | | | | | | | |
Db 149 AAACGCAAAAGAAAGCATTAAAGCTTAACTATGTAGAAGCAGTAGCTTTGATTAGTGCC 208
| | | | | | | | | | | | | | | | | |

QY 329 CATATTATGAACGAGCGCGCTGGGAAAACCGTTGCCAGCTTATGGAAGATGC 388
| | | | | | | | | | | | | | | | | |
Db 209 CATATTATGAACGAGCGAGCTGGTAAAGACTCGGCTGATTAATGATGCAAGAGGG 268
| | | | | | | | | | | | | | | | | |

QY 389 ATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGTGGGTAAATATGTTCCCGATCTA 448
| | | | | | | | | | | | | | | | | |
```

QY	1494	AGCTCTAGCAAAAAACAACTTGTAGAACAAAGTAGAAGCGGCGCGAGTTGGTTTTAAATTG	1553
DB	1406	GCTTCTAAACGATCGAGCTTAGCCGATCAAAATTGAAGCCGGTGGATGGCTTTGGCAATT	1465
QY	1554	CATGAAGACTGGGGCACAACACCAAGTGGCATCGATCACTGCTTTGAGCGGTGGCAGATGAA	1613
DB	1466	CACGAAGACTGGGGCACACACTCTTCTGCNATCAATCATCGGTTAGATGTTGGGACAAA	1525
QY	1614	TACGATGTCAAGTTTGTATCCACACCGATACAGTCAATGAGGCAGGTTATGTAGATGAC	1673
DB	1526	TACGATGTCAAGTTCGTATCGCCACAGACACTTTGAATGAAGCCGGTTGTGTAGAAGAC	1585
QY	1674	ACCCTAAATGCAATGAACGGGGCGGCATCCATGCCTACCAATGAGGAGCGGGTGA	1733
DB	1586	ACTATGGCTGCTANTGCTGGAAGCACTATGCACACTTTCCACACTGAAGCGCTGCGCGC	1645
QY	1734	GGACACTCACTGATGTTATCACCATGGCAGGCGAGCTCAATATCTTACCCCTCCTCCACC	1793
DB	1646	GGACAGCTCTGATATTAAGTAGCCGGTGAACAACAATCTTCCCGCTTCCACT	1705
QY	1794	ACCCCACTATTCCTATACCATTAATAACGGTTGCAGAACACTTAGACATGTCATGACA	1853
DB	1706	AAACCCACCATCCCTTTCCCGTGAATACAGAAGCAGACACATGGACATGCTTATGGTG	1765
QY	1854	TGCCACCACTAGACAAAACGATCCGGGAGGATTTACAAATTTCTCAAGCCGTATCCGC	1913
DB	1766	TGCCACCACTGGATAAAGCAATTAAGAAGATGTTCAAGTTCGCTGATTAAGGATCCCGC	1825
QY	1914	CCCGGCTCTATCCGGCTGAAGATGTCTCCATGATATGGGTGTGATCGCGATGACAAAGC	1973
DB	1826	CCTCAAAACCATTCGGCTGNAGACACTTTGCATGACATGGGGATTTTCTCAATCACCACT	1885
QY	1974	TCGGATTCGCAAGCAATGGGGGTGACGGCGAAGTGAATCCTCGAATTGGCAGACTGCG	2033
DB	1886	TCTGACTCTCAAGCGATGGCGGTGTGGGTGAAGTTATCACTAGAACTTTGGCAACAGCT	1945
QY	2034	GATAGAATAAAGAAATTTGGTGAAGCTTCCTGAGATGGGCAAGTAACGATAATTTTC	2093
DB	1946	GACAAAAACAAGAAATTTGGCCGCTTGAAAGAGAAAAAGCGCATACGCAAACTTC	2005
QY	2094	CGCATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTTGACCCACGCGTGAGC	2153
DB	2006	AGGATCAAAAGCTACTTGTCTAATAACACATTAACCCAGCGATCGCTCATGGGATAGC	2065
QY	2154	GAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTGTGFTGTGGAATCTCTGCC	2213
DB	2066	GAGTATGTAGTTTCAGTAGAAGTGGGCAAAAGTGGCTGACTTGGTATTTGTGGAGTCCAGCA	2125
QY	2214	TTTTTTGGGTAAAAACCAAAATCGTGATCAAAAGCGGTATGGTGTCTTCTCTGAAATG	2273
DB	2126	TTCTTTTGGCGTAAAAACCAACATGATCATCAAAAGGCGGATTCATTCGGTTAAGCCAAATG	2185
QY	2274	GGCGATCTTAACCGCTCTGTGCCCACTCCCCCAACCGGTTATTTACCCGCAAAATGTTTGGG	2333
DB	2186	GGCGATCGNAACGCTTCTATCCCTACCCCAACACCGGTTATTTACAGAGAAATGTTGCT	2245
QY	2334	CATCACGGAAGCGAAATTTGACAACGATCACTTTTGTGTTTCCAAAGTCGCTATGAA	2393
DB	2246	CATCATGGTAAAGCTAAATACGATGCAAAACATCACTTTTGTGCTCAAGCGGCTTATGAC	2305
QY	2394	AATGGCTGAAAGAAAGCTGGGCTTAGAGCGGCCAGTTCCTACGGTCAAAAACCTGCCGT	2453
DB	2306	AAAGGCATTAAGAAGAAATTAGGACTTTGAAGACAAAGTGTGCGGTAAAAAATTTGCAGA	2365
QY	2454	AACATCAAGAAAAAGACTTCAAGTTTCAACGACAAAACCGCAAAAAATCAACGTCGATCCG	2513
DB	2366	AATATCACTAAAAAGACATGCAATTCACGACACTACCGCTCACATTTGAATCAATCCT	2425
QY	2514	AAACCTTCGAGTCTTTGTAGATGGCAAACTCTGCACTCTTAAACCCACTCGCAAGTG	2573
DB	2426	GAACTTACCATGTGTTCTGGTAGGCAAGAAAGTAATCTTAAACACAGCAATAAAGTG	2485

```

Qy 2574 CCTCTAGCCAGCGGTACACTTCTTCTTAGG 2604
Db 2486 AGCTTGGCGCAACTCTTTAGCATTTTCTTAGG 2516

RESULT 8
US-07-732-242C-8
; Sequence 8, Application US/07732242C
; Patent No. 5298399
; GENERAL INFORMATION:
; APPLICANT: Uozumi, Takeshi; Masaki, Haruhiko;
; APPLICANT: Hidaka, Makoto; Nakamura, Akira;
; APPLICANT: Madao, Michihisa; Yoneta, Yasuo
; TITLE OF INVENTION: Gene of Urease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frischauf, Holtz, Goodman & Woodward, P.C.
; STREET: 600 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10016-2088
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.72mb
; COMPUTER: IBM PC compatible (NEC PC-9801 RX)
; OPERATING SYSTEM: MS DOS
; SOFTWARE: ASCII Form
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/732,242C
; FILING DATE: 19910718
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JPN 2-210178
; FILING DATE: 10-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Herdert
; REGISTRATION NUMBER: 17081
; REFERENCE/DOCKET NUMBER: 910532/HG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)972-1400
; TELEFAX: (212)370-1622
; TELEX: 236268
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6131 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-732-242C-8

Query Match 23.2%; Score 667.4; DB 2; Length 6131;
Best Local Similarity 59.0%; Pred. No. 1.3e-155;
Matches 1460; Conservative 0; Mismatches 951; Indels 63; Gaps 11;

Qy 185 ATTAAACACAGGAGTAATAGGTGAACTCACACCCAAAGACGACAGAAAAGTTCTTGTTA 244
Db 317 ACTTAAATAGGAGTTATCCATAAAGTCTTACGTTGAAATGGAAGCTCATGATT 376
Qy 245 TATTATGCGGGCGAAGTGGCTAGAAAGCGCAAGAGAGGGCTTTAAAGCTCAACCAACCC 304
Db 377 GTAGTGGCGGCTGACTTGGCCCGCGTCGTAAAGAGCGGGGCTTAAATTTAAATTTATCCT 436
Qy 305 GAAGCATTTGCTTACATTAGTGGCCCATATTATGGACGAGCGCGCTGCAAAAAAACCC 364
Db 437 GAAGCTGTGCAATGATTTACATTAAGTGTGAGAGCGGGCGGGATGG---AAAAACG 493
Qy 365 GTTGCCGAGCTTATGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCGGG 424
Db 494 GTAGCTCAGTTAATGCATACGTTGCAACGATTTCTTACAAAGAGATGTAATGGAAGG 553
Qy 425 GTGGGTAATATGTTCCCGATCTAGGTGTAGAACCCACTTTCTTGATGTGCAAACTT 484

```

Db 554 GTTCCGAAATGATCCCGGATATTTCAAATTGAGCGAACCTTTTCCTGATGGAACAAAGCTT 613
Qy 485 GTAACCTGTGAATTTGGCCCATC--GAACCGAGATGAGACTCTCAAAGCGGGCGAAGTGAAT 542
Db 614 GTACCGTTTCATGACCCGATCCGTTAAATGAGGAGACGTACGATGATACACAGGGAGTA 673
Qy 543 TTGGTTCCGATAAAGAC--ATCGAGCTCAATGACAGGCAAGAAGTAACCGAACTTGAGGT 600
Db 674 TGTATTAAAAAAGAACCTTATTTTATGCAATCAAAATAAGCAGACGATCAAGATTCGCGT 733
Qy 601 TACTAATGAAGGGGCTTAAATCTTTGCAATGTGGGTAGGCATTTCCACTTCTTTGAAGCTAA 660
Db 734 GTTAAACCGGGCGATCGACTGTTTCAGTTGGTTCCCATTTTCATTTTGAAGTGA 793
Qy 661 CAAGGCACTAAATTCGATCGTGAAGAAAGCTATGGAAGCCCTAGATATTCCTCTGG 720
Db 794 TCAATCGCTTCAATTTTCATCGTGAAGAAAGCATTTGGCATGCTTTGAATATTCGCGCTG 853
Qy 721 CAACCGCTAGCATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTCTGGTG 780
Db 854 AACGGCGTTTCGCTTCGAGCCCGAGATGCGAAAGATGAGAAATTAATTCATTTTCAGG 913
Qy 781 CAGTAAAAAAGTGAATTCGATGAACGGCTTGTGAATAACATCGCGATGAACGCCATA 840
Db 914 TGAACGCAAGTGTATG-----TTTAAATATGTAACGATG----- 951
Qy 841 ACATAAGCGCTTGACAAAGCGGAATCTCACGGATTTTCAAGTAAGGAGACTCCCATGA 900
Db 952 -----GATCAGTTTGAATGGGAAAGAAAATGAGTT-----TTTCGATGTC- 995
Qy 901 AAATGAAMAAACGAATATGTAAATACCTACGACCCCAAGCCCAAGGCGATAAAGTGGCT 960
Db 996 -----GAA-----AGCAATATGCGATATGTTTGACCAACTGTGCGGCGACCCATTCGTT 1046
Qy 961 TAGGATACCGATCTTTGGCGAGAGTAGAACAATGACTATACCCTATGCGCAAGAAC 1020
Db 1047 TGCAGATTCAGAAATGTTTATCGAAATTGAAATGGAAGGACTATACAACTGATGGAGTAGG 1106
Qy 1021 TTAATTTTGGCGGGTAAATCTATCGTGAGGTATGGGTTCAG-AGCAATAGCC--CTG 1077
Db 1107 TAAAGTTTGGCGGCGCAAGGTGATCCGAGATGGAATGGGCGAGCATCTTTGGCGCAA 1166
Qy 1078 ATGAAACACCCTAGATTTAGTTCATCACTAACCGCATGATATGCACTACACCGGATTT 1137
Db 1167 GCGATGAATGCGTCCGATCTCGTATTAAACAATGCGATTAATGTTGATTAACAGGATTT 1226
Qy 1138 ACAAAGCGCATTTGGATTTAAAAACGGCAAAATCCATGCAATGCGCAAGCGAGGAACA 1197
Db 1227 ATAAAGCAGATATCGGCATAAAGATGGAATGATTCCTCCNTAGGAAAGACGGGGAAAC 1286
Qy 1198 AGGACATGCAAGATGGCGTAAGCCCTCATATGTTGTTGGGTGGGCAAGACACTAG 1257
Db 1287 CGTTGTTAATGAGCGGGT-----CGATATGTTGATGAGCGAGCAAGATCAATAG 1340
Qy 1258 CAGGGGAAGGTATGATTTATPACCGCTGGGGGAATCGATTCACACACCCACTTCCTTCTC 1317
Db 1341 CGCAGAAGGGATGATTTGTGACAGCGGAGGAATAGATGCTCATATTCATTTTATTTGCC 1400
Qy 1318 CACAAATTCCTACCGCTCTAGCCATGCGGTTCACCAATGTTTGGAGGCGGCAAG 1377
Db 1401 CTCAGCAAAATCGAAACCGCTCTTTGCACTCGGGTGTGACCACTATGATTTGGGAGGAACAG 1460
Qy 1378 GTCCTGTAGATGCGAATGCGACTACTATCACTCGGGCAAAATGGAATTCGACCGCA 1437
Db 1461 GACCCGCTACAGGCACAAAATGCCACTACTTGTACACCGGGCGCTTGGAAATATCATCGTA 1520
Qy 1438 TGTTCGCGCAGCAGGAAGATTTCTATGAATGTGGGCTTTTTTGGCAAAAGGCAATAGCT 1497
Db 1521 TGCCTCAAGCAGCCGAAGAAATCCCGATAACTTGGGCTTTTTTAGGAAAGGGAACCTGTT 1580
Qy 1498 CTAGCAAAAAACAACTTTGTAGAAACAATGAGAGCGGCGCGATGTTGTTTAAATTCGATG 1557
Db 1581 CAGATGAGGCTCTCTTTAAAGGAACAAATTTGAAGCGGAGCGGTGGGATTAAGCTTCACG 1640

Qy 1558 AAGACTGGGCGACAACACCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATGAATACG 1617
Db 1641 AAGATTGGGGATTCGACGGCGCGCTATTGATACATGTTTGAAGTGGCGGATCGATATG 1700
Qy 1618 ATGTCAAAGTTGTTATCCACCCGATACAGTCAATGAGCGAGGTATGTAGATGACACCC 1677
Db 1701 ATGTCAAAGTAGCGATTCATACAGACATTTAAATGAAGCGGATTTGTTCGAGTACTT 1760
Qy 1678 TAAATGCAATGAACGGCGCGCATTCATCCCTACACATTTAGGGAGCGGGTGGAGGAC 1737
Db 1761 TGAAGCCCATAGACGGTGCAGTGAATTCATACCTATCATACAGAAAGGGCTGGCGGGGAC 1820
Qy 1738 ACTCACTGATGTTATCACCATGCGAGCGAGCTCAATATTCTACCTCTCCACACCC 1797
Db 1821 ATGCTCCGGATATTATAAAGCGCGCTTCCCGAATATTTTCCGCTTCTTCACGAATC 1880
Qy 1798 CCACATTTCCCTATACCATTAATACGTTTGAGAGACACATTTAGACATGCTCATCACATGCC 1857
Db 1881 CAATCGACCTTATACTATCAATATCTTTGGAAGAGCATTTAGATATGTTAATGGTTGCC 1940
Qy 1858 ACCACTAGACAAACGCGATCGCGAGGATTTACAAATTTTCTCAAAAGCCGTATCCGCCCG 1917
Db 1941 ACCACTAGACGCTAATATTCAGAGGATATTGCTTTTGCAGATTCCACGCATACGGAAG 2000
Qy 1918 GCTCTATCGCGCTGAAGATGTCTCATGATATGGGTGTGATCGCGATGACAAGCTCGG 1977
Db 2001 AGACCATCGCGCGAAGATGTTTACATGATTTAGGCGTTTTTCAGCATGATTTCTGCTG 2060
Qy 1978 ATTCGACGCAATGGGCGTGCAGCGAAGTATCTCGAATTTGGGACACTTGGCAGCTGCGGATA 2037
Db 2061 ATTCACAGCGATGGGCGGAGTAGGAGAGTATCATTCGTACGTGGCAACGGCTGACA 2120
Qy 2038 AGAATAAAAAGAAATTTGTTAGTACGTTCTCTGAAGATGGCAAGATAACGATAATTTCCGCA 2097
Db 2121 AGATGAAAAAGCAAGAGAGGAGTTTACAGAGACAAATGGTGTGGAGACAACATTTCTGTG 2180
Qy 2098 TTAAGCGCTCATCTTCCAAATACATATCAACCCCGCTTTGACCCACGCGGTGAGCGAGT 2157
Db 2181 TGAAGCGTTATATTGSCAAATATACGATCAATCCGCGCATTCCTCATGTTATTCGGAT 2240
Qy 2158 ATATCGGCTCTGTGGAAGAGGCAAGATCGCGCATCTGGTGTGTGGAATCTCGCCTTTT 2217
Db 2241 ATGTGGGTCTCTGTTGAAGTGGGAAATTTAGCTGATTTAGTGTGTGGAATCTCGCTTTT 2300
Qy 2218 TTGGCGTAAACCCCAAAATCGTGAATCAAGGCGGTATGTTGCTTCTCTGAAATGGCGG 2277
Db 2301 TTGGTGTGAACCTGACTGGTCTTTAAAGGAGGAATGATTCCTTACGACACTATGGAG 2360
Qy 2278 ATTCTTAACGCGCTCTGTGCCCATCTCCCAACCGGTTTTTATACCGCGAAATGTTTGGGCATC 2337
Db 2361 ATCCCAATGCCAGCATTCGGACACCGCAGCGGTTTTTATATCGTCCGATGTTTGCAGCGA 2420
Qy 2338 ACGGCAAGCGGAAATTTGACACCGAGATCACTTTGTTGTTTCCAAAGTCGCTATGAAAAATG 2397
Db 2421 AAGGAGATGCCAAATATCAAAACGCTATACACCTTTGTTTTCGAAAGCAGCGTATGAAAAAG 2480
Qy 2398 GCGTGAAGAAAGAAAGCTGGGCTTAGAGCGCCCAAGTTCTACCGGTCAAAACCTGCGTAACA 2457
Db 2481 GCATTCATGAACAGTTGGGTTTGAAGAAAAAGGTGAACCGAGTCCATGGAAATTCGAAAT 2540
Qy 2458 TCACCAAGAAAGACTTCAAGTTTCAACGACAAAAACGGCAAAAAATCACCGTCGATCCGAAAA 2517
Db 2541 TGACGAAAAAGATTTAATTTTGAACGATAAAAACCCCAAAATTTGACGTCGATCTCTCAGA 2600
Qy 2518 CTTTCGAGGTCTTTGTAGATGCGAAACCTCTGACCTCTAAACCCACCTGCGCAAGTGCCTC 2577
Db 2601 CATATGAAGTAAAGTAGACGGTCAATTTAGTGACATGTGAACCGGACAGAAATCGTCCCTA 2660
Qy 2578 TAGCCGAGCGCTACACTTTCTTAGGCACAATGCCCCCTTTGGGGGAGGTTATTTTAG 2637
Db 2661 TGGCACACCGGATTTTCTTATTTTGGGTGAGAAAAACATGATGTTGAAAAAGTAGTCG 2720

QY 1436 CATGTTGGCGCAGCAGAGAGTATTTCTATGAAATGTGGCTTTTTTGGGCAAGGCAATAG 1495
 Db 563522 TATGTTTCAAGCGCGAGAGGCTTCCCGGTAAACGTGCGATTTTTTGTGAAGGCACTG 563563
 QY 1496 CTCTAGCAAAAACAACTTGTAGAACAGTGAAGCGGCGCGATGTTGTTTAAATTGCA 1555
 Db 563562 TTCAACCTAGATCTCTGCGTGAGCAAAATGAAGCGGTGCATTAGGTTTAAAAATCCA 563503
 QY 1556 TGAAGACTGGGGCACAACCAAGTGCGATCGATCACTGTTGAGCGGTGGCAGATGAATA 1615
 Db 563502 CGAAGACTGGGGTGCAACGCCCTGCGGTGATTTGCTTCTTAAAGTAGCAGATGAAT 563443
 QY 1616 CGATGTGCAAGTTTGTATTCACACCGATACAGTCAATGAGCGAGGTTATGTAGTAGCAC 1675
 Db 563442 GGAATTCAGGTGGCCATTACACAGACACCGCTAAATGAAGTGGCTTTTGGGAAGCAC 563383
 QY 1676 CCTAAATGCAATGAACGGCGCGCCATCCATGCTTACACATTCAGGAGCGGTGGAGG 1735
 Db 563382 GATGAAGCGATTGATGAGCGAGTCAATTCATCTTCCATACGAGGCGGCGAGTGGTG 563323
 QY 1736 ACATCTACCTGATGTTATCACCATGAGCGAGGAGCTCAATATTTACCTCTCCACCCAC 1795
 Db 563322 TCATGCACCTGACATCATTAAGCAGCGATGTATTTCAACGATTTACCTGCTTCAACCA 563263
 QY 1796 CCCCACTATTCCTTATACCATTAATACGTTGAGAGCACTTAGACATGCTCATGACATG 1855
 Db 563262 CCCGACTCGTCCGTTTACCAAAAACACCATTTGATGAAATTTGGATATGTTGATGGTTG 563203
 QY 1856 CCACCACTAGACAAACGATCCGCGAGGATTTACAATTTTCTCAAGCGGTATCCGCC 1915
 Db 563202 CCATCACTTAGATTAACGCGTCCGGAAGAGTATGCTTTTCCGATAGCGGTATCCGCC 563143
 QY 1916 CGGCTCTATCGCGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATGACAAAGTCT 1975
 Db 563142 TGAACCAATTCGACGAGAGATATTTTGCATGATATGCGGTCTTCTCCATTAATGAGTTC 563083
 QY 1976 GGATTCGCAAGCAATGGGCGTGACGCGAGTATCTTCGAACTTTGGCGAGACTCGGGA 2035
 Db 563082 AGACTCTCAAGCGATGGGAGTATTTGGCGAAGTGTATTCGTACATGCGCAAACTCGAGA 563023
 QY 2036 TAAGATAAAAAGAAATTTGTAAGCTTCTGAGATGGCAAGATGCAAGATATTTCCG 2095
 Db 563022 TAAGATGAATGCAAGCTGTGAGCT-----AGTAAATGAAGGAACCAATATCTTCCG 562969
 QY 2096 CATTAAAGCGCTACATCTCCAAATACATCAATCAACCCCGCTTTGACCCACGCGGTGAGCGA 2155
 Db 562968 TATTAAAGATATATCGGAATACACCATCAACCCAGCAATTCGACATGGTATGCGGA 562909
 QY 2156 GTATATCGGCTCTGTGGAAGAGGCAAGATGCGGACTTGGTGGTGTGGAAATCTGCGCTT 2215
 Db 562908 GCATATGGCTCGTTAGAGTGGTAAATCGCAGATATCGTGTATGGAACCGATGTT 562849
 QY 2216 TTTTGGGTTAAACCCCAAAATCGTATCAAGGCGGTATGGTGTCTTCTGAAATGGG 2275
 Db 562848 CTTTGGGTTAAACCTGAAGTGGTATTAAGAAAGGCTTTATTAGCTATCGCAAAATGGG 562789
 QY 2276 CGATTTCAAGCGCTGTGCGCACTCCCAACCGTTTATACCGCGAAATGTTGGGCA 2335
 Db 562788 CGATCCAAATGCTCAATTCACACCGCAACTGTATTTACCGTCCAAATGACGTGC 562729
 QY 2336 TCACGCAAGGCGAAATTTGACACCGCAATCACTTTTGTTCCAAAGTCGCTATGAAA 2395
 Db 562728 ACAAGCTTAGCAACCGCAACACAGCAGTATTTTGTGTTTCAAGCGCTGAAAAGC 562669
 QY 2396 TGGCGTGAAGAAAGCTGGGCTTAGAGCGCCAAAGTTCTACCGGTCAAAAACCTGCCGTAA 2455
 Db 562668 TGAATATTCGTCGAAGTTCGGTTTACAAAGAAACCAATGCTGTGTAAGAGCTGCCCAA 562609
 QY 2456 CATCAACAGAAAGACTTCAAGTTCAACGACAAACGCGAAATACCCCTCGATCCGAA 2515
 Db 562608 CGTAGATAAAAAGATCTGGTTCATATGATGTAAACACCAACATTTACTGTTGATGCTGA 562549
 QY 2516 AACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACTCTTAAACCCACCTTCGCAAGTGCC 2575

Db 562548 ACCTTATGAGTTCAGTGCAGCGAGAGTAAATTAACCTGTGAACCAAGTGGATAGCGTACC 562489
 QY 2576 TCTAGCCCGCGCTACACTTTCTTCTTA 2602
 Db 562488 ATTGGTCAAGCATATTTCTTCTTCTTA 562462

RESULT 10
 US-09-643-990A-1/c
 ; Sequence 1, Application US/09643990A
 ; Patent No. 6528289
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert D. Fleischmann
 ; Mark D. Adams
 ; Owen White
 ; Hamilton O. Smith
 ; J. Craig Venter
 ; TITLE OF INVENTION: The Nucleotide sequence of
 ; the Haemophilus influenzae Rd Genome, Fragments
 ; Thereof, and Uses Thereof
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville,
 ; STATE: MD
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3 1/2 inch diskette
 ; COMPUTER: Dell Pentium
 ; OPERATING SYSTEM: MS DOS V6.22
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/643,990A
 ; FILING DATE: 23-Aug-2000
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/487,429
 ; FILING DATE: 1995-06-07
 ; APPLICATION NUMBER: 08/426,787
 ; FILING DATE: 1995-04-21
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kenley K. Hoover
 ; REGISTRATION NUMBER: 40,302
 ; REFERENCE/DOCKET NUMBER: PB186PLC1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 301-610-5790
 ; TELEFAX: 310-309-8439
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1830121 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 ; US-09-643-990A-1

Query Match 23.1%; Score 665.4; DB 3; Length 1830121;
 Best Local Similarity 57.0%; Pred. No. 5.9e-154;
 Matches 1452; Conservative 1; Mismatches 1057; Indels 37; Gaps 14;

QY 83 TTTTATAATAATTACTTATATCATATATAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 142
 Db 564998 TATATAAGGAGGATTTTTTTTCTTTTCAAGGGCTTTACAAAAATTTCTGACAAAATATTTT 564939
 QY 143 AAAAAGTAAAGAAATAGGACTATAATCCCATTCCTTTTAAATTTTAAACACAGGAGT-A 201
 Db 564938 TCCTATTTTAGATTTTTCGTAAATTCGTCGATTCGTTATTAATAATTAAGGATGA 564879
 QY 202 ATAGGTGAAACTCACCCCAAGCAAGAAAGTCTTGTATATTAATGTCGGGCGCAAGT 261

Db 564878 AAATATGCACTTAACCTCCAGAGAACAGAAAAAAGTGAATGCTTTTCTCGCGGGCAACT 564819
Qy 262 GGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGCAAGCCATTTGCTTACAT 321
Db 564818 TCGCGCAAAACGCAAGCAGCGCGGTAAATTAACCTATCCAGAAACTATTTGCTTATAT 564759
Qy 322 TAGTGCCCATATATGACGAAAGCGCGCTGGGAAAAAACCCTTGCCAGCTTATGGA 381
Db 564758 TGCTAGTCAATTTACAAGAGGCGCAGAGAAAGGAA---TGAGTGTAGCGGAATCATGCA 564702
Qy 382 AGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAAATGCCCGGGTGGGTAAATATGCTTCC 441
Db 564701 ATATGGCGCAACACTTTTAACCGTTGATGATCATGGAAGGTGGCGGAAATGGTTCA 564642
Qy 442 CGATCTAGGTGTAGAGCCACTTTCTCTGTATGTTACGAAACTTGTAACTGTGAATTTGCC 501
Db 564641 TGAAGTCCAGATTGAAGCTACTTTTCCCGATGGCAGCAAACTTGTACCGTGCATAATCC 564582
Qy 502 CATC-GA-ACCAGATGACACTTCAAGCGGGCGAAGTGAATTTGG--TTGCGATAAAG 557
Db 564581 AATCAGATAAACCGWAGGGTGGGCTTTAGCCCAACAAAATAAAAATATCAATGTTGGGCTAA 564522
Qy 558 ACATCGAGCTCAATGCGAGCAAGAGTAACCGAACTTTGAGGTACTAATGAAGGGCCTA 617
Db 564521 AGCCACCTTACAAGGAAACAAAGATATGATCCAGCGGATACCAATTAGCGGAAGCGGA 564462
Qy 618 AATCCTTTGC--ATGTGGGTAGCCATTTTCCACTTCTTTGAAGCTAACAGGCACATAAAAT 675
Db 564461 TATTTCTCGCTAATGTGCGCAGAAAAACGTAATAAATCGAAGTAACAAATTCAGGCGACCG 564402
Qy 676 CGATCGTGAAAAAGCT---ATGGCAACGCTTGAATATTCCTCTGGCAACACGCTA-- 730
Db 564401 CCCAATTCAGTGTGGCTCGCATTAACCAATTTTGTGAACCAATAATGCGCTTAAATTTGA 564342
Qy 731 -CGCATTTGGGCGAGCAACAAACCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAA 789
Db 564341 CCGCACTTTGGCAGCGTGAATTCGCGCTTAAATGTTCCATCTCGCAATGCGGTGCTTTGA 564282
Qy 790 AGTGATTGGCATGAACGGGCTTGTGAATAACATCGCGATGAACGCCATAAACATAAAGC 849
Db 564281 ACCCGTGAAGTGAATCAGTGGAAATGATGTGCTTTTGGTGAACCA-ATCATTTATG 564223
Qy 850 GCTTGACAAGCGCAAACTCTCAGGATTTATCAAGTAAG--AGACTCCCATGAAAAATGAA 907
Db 564222 GTTTCCATAATCAAAATGATGCGAAATTTAAGGTAGGCGAAGTGGCATTAACAAATTC 564163
Qy 908 AAAACAAGATATGTAAATACCTACGAGCCCAACCAAGCGGATAAAGTGGCTTAGGAGA 967
Db 564162 AAGAGCGCAATATGTAGCAACTTATGCTTCCAAACAGTTGGCGATAAAGTCCGTTTAGGCGA 564103
Qy 968 TACCGATCTTTGGGCGAAGTAGAACATGACTATACCACTTATGCGGAAGCACTTAAATTT 1027
Db 564102 TACCAATTTATGGGCAACCATTTGAACAGATTTATTTGACCAAGGTGATGATGTAAATTT 564043
Qy 1028 TGGCGCGGGTAAAACTATPCGTTGAGGGTATGGGTGAGAGCAATAGCCCTGATGAAACAC 1087
Db 564042 TGGTGGCGGTAAAGCGTGGTGTATGCTCAAGAGGTACGGCAACTCGCGACA 563983
Qy 1088 CC-----TAGATTTAGTCATCATTAACGCGATGATATGGA--CTACA-CCGGGAT 1135
Db 563982 TCCGAATGTATTTGGATTTTGTGATTTACCAACGTTGATGATCATTTGATGTCTTAAATTTAGGCAT 563923
Qy 1136 TTACAAAGCCACATTTGGGATTTAAAGCGCAAAATCCATGCAATTCGACAGGAGGAAA 1195
Db 563922 TATCAAGCCCATATTTGATTTTCTGTATGGGCGTATTTGGGTATTTGGACAGAGGTAA 563863
Qy 1196 CAAGCATGCAAGATGGCGCTTAAAGCCCTCATATGTTGTTGGGTGGGCGACAGAGCACT 1255
Db 563862 CCCTGACACCATGATTAACGTCACACCAAAATATGATTTATCGGTGAGCAACGGAATTTCA 563803
Qy 1256 AGCAGGGGAAGGTATGATTTATTCGCTGGGGGAATGCAATTCACACACCACTTCCTTTC 1315
Db 563802 TAAACGCTGCACATTTAAATTTGCAACCGCTGGTATTCGATATCCCATCCCACTTCACTTTATTTG 563743

Qy 1316 TCCACAAACAATTTCCCTA CCGCTCTAGCCCAATGGCGTTTACAAACATGTTTGGAGCGGCAC 1375
Db 563742 TCCACAAACAAGCACCAATGCAATTTCAAAGTGGCGTTACCAAGTAAATTTGGTGGTGAAC 563683
Qy 1376 AGGTCTCTAGATGGCAGAAATGGAGCTACTATCACTCGGCGCAATGGAATCTTGACCG 1435
Db 563682 TGGCCCTCTGATGTGTACACAGCAACCACTTGTATCCCTGGCGCATGTGTATATGGAACG 563623
Qy 1436 CATGTTGGCGCAGCAGAGAGATTTCTATGAATGTGGCTTTTTCGGCAAGCGCAATAG 1495
Db 563622 TATGTTTCAAGCGGCGAAGCCCTTGCCTGTAACGTCGATTTTGGTAAAGGCAACTG 563563
Qy 1496 CTCTAGCAAAAAACAACCTTGTAGAACAGTAGAAGCGGCGGATTTGGTTTAAATTTGCA 1555
Db 563562 TTCAACCCCTAGATCTCTCTGCTGAGCAAAATGAAGCGGCTGCAATAGGTTTAAATAATCCA 563503
Qy 1556 TGAAGACTGGGCGACAAACCAAGTGCATCACTCTTGAAGCGTGGCAGATGAATA 1615
Db 563502 CGAAGACTGGGGTGCACAGCCCTGCGGTGATTTGCTTGAAGATGACAGATGAAT 563443
Qy 1616 CGATGTCAAGTTTCTGTATCCACACGATACAGTCAATGAGGCGAGGTTATGTAGATGACAC 1675
Db 563442 GGAATTTCAAGTGGCCATTTCAACAGACACGCTAAATGAAGTGGCTTTTGGAAACAC 563383
Qy 1676 CCTAAATGCAATGAACGGGCGCCCATCCATGCCCTACCACTTGAAGGAGCGGTTGAGG 1735
Db 563382 GATGAAGCGATTTGATGAGCAGTCAATTCATCTTTCCATACGAGGCGGCGAGTGGTG 563323
Qy 1736 ACATCACTGATGTATCACCATGGCAGGCGAGCTCAATATTTCTACCTCTCCACCAC 1795
Db 563322 TCATGCACCTGACATCATTTAAAGCAGCGATGATTTCAACAGTATTTACCTGCTTCAACCAA 563263
Qy 1796 CCCCACTATTCCTATACCATTAATACGTTGAGAACTTAGACATGCTCATGACATG 1855
Db 563262 CCGGACTGTCGTTTACCAAAAAACCATTTGATGAACATTTGGATTTGGATTTGTTG 563203
Qy 1856 CCACCACTTAGACAAAACGATCCGCGAGGATTTACAATTTTCTCAAAGCGGTATCGCCC 1915
Db 563202 CCATCACTTAGATAAACGCGTCCCGAGACGCTAGCTTTTCCGATAGCGGTATCCGCC 563143
Qy 1916 CCGCTCTATCGGCGCTGAAGATGTCTCATGATATGGGTGTGATCGGATGACAGCTC 1975
Db 563142 TGAACCACTATGACAGAGATATTTTGCATGATATGGCGCTTCTTCCATTTATGATTC 563083
Qy 1976 GGATTCGCAAGCAATGGGCGCTGAGGAGGATTTCTCGAACTTTGGCAGACTCGGA 2035
Db 563082 AGACTCTCAAGCGATGGGACGTTTGGCGAAGTCTGTTATTCGTACATGGCAAACTGAGA 563023
Qy 2036 TAAAGATAAAAAAGAAATTTGGTAAGCTTCTCTGAAGATGGCAAGATAAACGATTAATTTCCG 2095
Db 563022 TAAAGATGAATTAACGCTGGTGGT-----AGGTAATGAAGGAAACGATAACTTTCCG 562969
Qy 2096 CATTAAGCGCTACATCTCCAAATACATATCAACCCCGCTTTGACCCACGCGCTGAGCGA 2155
Db 562968 TATTAACCGATATATCGGAAATACACCAACCCAGCAATTCACATGTTATTTGCGGA 562909
Qy 2156 GTATATCGCTCTGTGGAAGGCGAAGATCCCGACTTTGGTGGTGGGAATCTGCTT 2215
Db 562908 GCATATGGCTCGTTAGAAATGGGTAAATTCGAGATATCGTGTATTTGGAACCGATGTT 562849
Qy 2216 TTTTGGCGTAAAAACCAAAATCGTATCAAAAGCGGTATGGTGTCTTCTCTGAAATTTGGG 2275
Db 562848 CTTTGGCGTAAAACTGAAGTGTGATTAATAAGGCTTTATAGCTATGCGAAATTTGG 562789
Qy 2276 CGATTTCAACGCTGTGTCCTCCCAACCGGTTTATTTACCGGAAATGTTTGGGCA 2335
Db 562788 CGATCCAAATGCTCAATTTCCAAACCGCAACCTGATTTCTACCGTCCAATGTACGCTG 562729
Qy 2336 TCACGGCAAGGGAATTTGACACAGCATCACTTTTGTGTTTCCAAAGTCCGCTTATGAAA 2395
Db 562728 ACAAGGCTTAGCAACCGCAACCAACAGCAGTATTTCTTTGTTTCAAGCGCGCTGAAAAGC 562669

QY 2396 TGGCTGAAGAAAGAGCTTACGGCCAGCTTCTACGGTCAAAAAGTCCGCTAA 2455
Db 56268 TGATATTCGTCGGAAGTTCGGTTTACAAAGAAACCAATGCTGTGAAGGCTGCCGAA 56269
QY 2456 CATCACCAAGAAAGACTTCAAGTTCAACGACAAACGGCAAAAATCACCGTCGATCCGAA 2515
Db 56268 CGTAGGTAAGAAAGATCTGGTTTCAATGATGATACACCAACATTACTGTTGATGCTGA 562549
QY 2516 AACCTTCAGGCTTTTGTAGATGGCAAACTCTGCACCTCTTAAACCACTCGCAAGTCCC 2575
Db 562548 ACGTATTAAGCTTCGAGTGGACGGAGGTTAATTACCTGTGAACCAAGTGGATAGCGTACC 562489
QY 2576 TCTAGCCAGCGCTACACTTCTCTCTA 2602
Db 562488 ATTGGTCAGCGATATTTCTCTATCTA 562462

RESULT 11

US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; Patent No. 6846651
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

```
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature

Query Match      23.1%; Score 665.4; DB 3; Length 1830121;
Best Local Similarity 57.0%; Pred. No. 5.9e-154;
Matches 1452; Conservative 1; Mismatches 1057; Indels 37; Gaps 14;

QY      83 TTTTAAATAATTACTTATATCATATATAATAATAATATTACTTATATATAAAGTTAAT 142
DB      564998 TATATAAGGAGGATTTTTTTCTTCAAGGGCTTACAAAAAATCTGACAAATATTTT 564939

QY      143 AAAAAGTAACGAAATAGGACTATATATCCATTCGCTTTAAATAATTTAAACACAGGAGT-A 201
DB      564938 TCCTATTTTAGATTTTACGTAAATTCGCTGACTTCCTGATATAATAATAAGGATGA 564879

QY      202 ATAGGTGAACTCACCCAAAGAGAGAAAAGTTCTTGTATATATATATGCGGCGCAAGT 261
DB      564878 AAATATGCACCTTAACTTCCAGAGAACACAGAAAAAATGATGCTTTTCTCGCGGCGCAACT 564819

QY      262 GGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGGAGCCATTGCTTACAT 321
DB      564818 TCGGCAAAACGCAAGCAGCGCGCTTAAATTTAAACTATCCAGAAACTATTGCTTATAT 564759

QY      322 TAGTGCCCATATATATGGACGAAGCGCGTGGAAAAAACCCTTGCCCGAGCTTATGGA 381
DB      564758 TGCTAGTCATTTACAAGAGGCGAGCAAGAGAAGAA---TGAGTGTAGCGGAAGTCAATGCA 564702

QY      382 AGAGTGCATGCACTTTTGTAAAAAGATGAAGTAATGCCCGGGTGGTAAATATGTTCC 441
DB      564701 ATATGGCGCAACACTTTTAAACCGTTGATGATGTCATGGAAGGTGTGGCGGAAATGTTCA 564642

QY      442 CGATCTAGGTGTAGAAAGCCACCTTTCCTGATGTCAGCAAACTTGTAACTGTGAATGGGCC 501
DB      564641 TGAAGTCCAGATTGAAGCTACTTTCCCGATGGCAGAACTTGTACCGTCATATATCC 564582

QY      502 CATC-GA-ACCAGATGAGCACTTCAAAGCGGCGAAGTCAAAATTTGG---TTGCGATAAAG 557
DB      564581 AATCAGATAACCGWAGGGTGGGCTTTAGCCCAAAATAAATAATCAATGTGTGGGCTAA 564522

QY      558 ACATCGAGCTCAATGCAGGCAAAAGTAACCGCAACTTGAGGTTACTAATGAAGGCGCTTA 617
DB      564521 AGCCCACTTACAAGCAACAAAGATATGATCCAGGCGAATACCAATTAGCCGAGGCGA 564462

QY      618 AATCCTTGC--ATGTGGGTAGCCATTTCCACTTTTGAAGCTAACCAAGGCACATAAATT 675
DB      564461 TATTCTCGCTAATGTGCGCAGAAAAACCGTAAAAATCGAAGTAACAAATTCAGCGCACCG 564402

QY      676 CGATCGTGAARAAGCCT---ATGGCAAAACGCTAGATATTCCTCTCGGCAACACGCTA-- 730
DB      564401 CCCAATTCAAAGTTGGCTCGCATTACCATTTTTTGAACCAATAATGCCCTTAAATTTGA 564342

QY      731 -CGCATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCTCTTTGGTGGCAGTAAAAA 789
DB      564341 CGGCACCTTTGGCAGCGTGAATCGGCTTAAATGCTCATCTGGCAATGCGGTGCGTTTGA 564282

QY      790 AGTGATTTGGCATGAACGGGCTTTGGAATAACATCGCGGATGAACGCCATATAACATAAAGC 849
DB      564281 ACCCGTGAAGTGAATAACAGTGAATAATAGTGTGTTAGTGTGTTAACCA-AATCATTTATG 564223

QY      850 GCTTGACAAGGCGAAATCTCAGGATTTATCAAGTAAGG--AGACTCCCATGAATAAGAA 907
DB      564222 GTTTCATAATCAAATTTGATGGCAAAATATAGGTAGGGCAAGATGGCAATTAACAATTTTC 564163

QY      908 AAAACAAGAAATATGTAATAACCTACGACCCCAAAAGGCGATATAAGTCGCTTAGGAGA 967
DB      564162 AAGAGGCAATATGTAGCAACTTATGGTCCACAGTTGGCGATAAAGTCCGTTTAGGCGA 564103

QY      968 TACCGATCTTTGGGCGAGAGTAGAACAATGACTATACCACTATGCGGAGAAAGCTTAAATT 1027
DB      564102 TACCAATTTATGGGCAACCATTTGAACAAGATTTATTGACAAAGGTGATGAGTGTAAATT 564043

QY      1028 TGGCGGGGTAAAACTATCCGTGAGGGTATGGTCAAGCAATATAGCCCTGATGAAACAC 1087
DB      564042 TGGTGGCGGTAAAGCGTCGCTGATGGTCTCAAGCGGTACGGCAACTCGCGCAAA 563983
```

Qy	1088	CC-----TAGATTAGTCATCATAACGCGATGATTATCGA--CTACA--CGGGAT	1135
Db	563982	TCCGAATGTTATGGGATTTTGTGATTACCAACGCGATGATCATTTGATGCTAAATTAGGCAT	563923
Qy	1136	TTACAAAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGCCATTGGCAAGCGCAGGAAA	1195
Db	563922	TATCAAGCCGATATTGGTATTTCGTGATGGGCGTATTGTGGGTATTGGCAACGAGGTAA	563863
Qy	1196	CAAGCATGCAAGATGCGCTAAGCCCTCATATGCTGTGGGTGTGGGCACAGAAGCACT	1255
Db	563862	CCCTGACACCATGGATAGCTCACACCAATATGATTATCGGTGCAAGCAGCAAGTTCA	563803
Qy	1256	AGCAGGGAAAGGTATGATTATTACCGCTGGGGGAATTCGATTACACACCCACTTCTCTTTC	1315
Db	563802	TAAACGGTGCATTTAAATTGCAACCGCTGGTATCCGATACCCACATTCATCTTTATTG	563743
Qy	1316	TCCACAAATTCCTACCGCTCTAGCCAAATGGCGTTACAAACCATTTTGGAGCGGCAC	1375
Db	563742	TCCCAACAAAGCAACAATGCAATTTGAAGTGGCGTTACCAAGTTAAATTTGGTGTGGAAC	563683
Qy	1376	AGGTCTGTAGATGGCAGAAATCGCACTACTCACTCGGGCAATGGAATCTGCACCG	1435
Db	563682	TGGCCCTGCTGATGGTACACAGCAACCACTTTGTACCTTGGCGATGGTATATGGAAG	563623
Qy	1436	CATGTCGCGCAGCAGAAAGATTTCTATGAATGTGGCTTTTGGCAAGGCAATGAG	1495
Db	563622	TATGTTTCAAGCGCAGAAAGCCTTGCCTGCTAAACGTCGGATTTTGTGTAAGGCACTG	563563
Qy	1496	CTCTAGCAAAAAACAACTTTGTAGAACATGTAAGCGGGCGGATTTGTTTAAATTGCA	1555
Db	563562	TTCAACCTTAGATCCTCTGCGTGAGCAAAATTTGAAGCGGGTGCATTAGGTTTAAAAAATCCA	563503
Qy	1556	TGAAGACTGGGGCAACACCAAGTGGCATCGATCACTGCTTGTAGCGTGGCAGATGAATA	1615
Db	563502	CGAAGACTGGGGTCAACGCGCTGCGTGATTGATTCTGCCTTTAAAGTACGATGAAT	563443
Qy	1616	CGATGTGCAAGTTTGTATFCCACCCGATACAGTCAATGAGCGAGGTATGTAGATGACAC	1675
Db	563442	GGATTTCAAGTGGCCATTCACACAGACACGCTAAATGAAAGTGGCTTTTGGGAAGCAC	563383
Qy	1676	CCTAAATGCAATGAACGGGCGCGCATCCATGCCATACACATTTGAGGGAGCGGTGGAGG	1735
Db	563382	GATGAAGCGATTGATGGACGAGTCATTCACTTCCATTCGGAGGGCGCAGGTGGTGG	563323
Qy	1736	ACATCACTGATGTATTATCCATGCGAGCGAGCTCAATATTCTACCTCCTCCACCAAC	1795
Db	563322	TCATGCACTTGACATCAATTAAAGCAGCATGATTTCAAACGATTACTGCTTCAACCA	563263
Qy	1796	CCCACATTTCCCTATACCATTAAATACGGTTTGCAGAACATTAGACATGCTCATGACATG	1855
Db	563262	CCCGACTCGTCGGTTTACCRAAAACACCATTTGATGAACATTTGGATATGTTGATGTTTG	563203
Qy	1856	CGACCCTTACAAACGATCCGGAGGATTTACAAATTTTCTCAAGCCGATTCGCGCC	1915
Db	563202	CCATCACTTAGATAAAGCGTGCAGGAGACGTAGCTTTTCCGATAGCCGATTCGCGCC	563143
Qy	1916	CGGCTCTATCGCGCTGAAGATGTCTCATGATATGGGTGTGATCGGATGACAAGCTC	1975
Db	563142	TGAACCATTTGCAGCAAGATATTTTGCATGATATGGGCGTCTTCTCATTATGAGTTC	563083
Qy	1976	GGATTGCAAGCAATGGGGCGTGAGCGCAAGTGAATTCCTCGAACTTTGGCAGACTGCGGA	2035
Db	563082	AGACTCTCAAGCGATGGGACGTAATTGGCGAAGTCGTTATTCTGTACATGGCAAACTGCAGA	563023
Qy	2036	TAAAGATAAAAAAGAAATTTGGTAAAGCTTCTTGAAAGATGGCAAGATTAACGATAATTTCCG	2095
Db	563022	TAAAGATGAATATGCAACGTTGGTGTAGCT-----AGGTAATGAAGGAACGATAACTTCCG	562969
Qy	2096	CATTTAAGCGCTACATCTTCCAAATACATCAACCCCGCTTTTGACCAACCGCGTGGCGGA	2155
Db	562968	TATTTAAACGATATATCGGAAATACACATCAACCCAGCAATTCGACATGGTATTTGCGGA	562909
Qy	2156	GTATATCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTGGTGTGTGGAATCTGCTT	2215

[illegible]

```

RESULT 12
US-09-543-681A-1857
; Sequence 1857, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1857
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-09-543-681A-1857

```

Query Match 22.5%; Score 648.2; DB 3; Length 1710;
Best Local Similarity 62.2%; Pred. No. 4.4e-151;
Matches 1052; Conservative 0; Mismatches 633; Indels 6; Gaps 2;

912	Qy	CAAGAAATATGTAATACTTACCGACCCACAAGCGGATAAAGTGCCTTAGGAGATACC	971
25	Db	CAAGCTTATCGCGATATGTTGGCCCAACAACAGGCGATCGTTTGGCATTAGCAGATACC	84
972	Qy	GATCTTTGGCGAGAAGTGAACAATGACTATACACACTATCGCGAAGAACTTAAATTTGGC	1031
85	Db	GAGCTGTTCTTGAATTTGAAAAGATTTCACCACTATCGCGAAGAGTCAAAATTTGGT	144
1032	Qy	GCGGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGATGAACAACCCCTA	1091
145	Db	GGTGGTAAAGTTATTTCGTGATGTTATGGGCCAAAGCCAAAGTTGTAGTCTGAGTGTGC	204
1092	Qy	GATTTAGTCACTCACTAAACCGGATGATTTCGACTACACCGGGATTTACAAGCCGACATT	1151
205	Db	GATGTTCTGATCACCAGTCGCAATTATTTTAGATTTATGGGGCATTTGAAAGCGAGATATT	264
1152	Qy	GGGATTTAAAAACGGCAAAATTCATGTCATTTGGCAAGGCAGGAAACAGGCAATGCAAGAT	1211

Db 265 GGCATTAAAGATGCCCGTATTTGTCGGTATTTGGCAAGCAGGTAATCCAGATGTTCA----- 320
Qy 1212 GGCCTAAGACCCCTCATATGCTGCTGGGTGTTGGGCAAGAGCACTAGCAGGCGGAAGGTATG 1271
Db 321 -GCCCAA-TGTGGATATTTGTCATTTGGCCCCGGAACAGAGTTGTTGCGGGAAGGTAAA 378
Qy 1272 ATTATTAACCGCTGGGGAATGATTTCAACACCACTTCCTTTCTCCACAACAATTCCT 1331
Db 379 ATTGTCACTGCTGGTGTATTTGATACCATATCCACTTTTATTTGTCGCAACAAGCCCAA 438
Qy 1332 ACCGCTTAGCCATGGGTATACACCACTTTGGAGCGGCAAGTTCCTGTAGATGGC 1391
Db 439 GAAGGTCTCGTTTCTGGGTAAACCACTTTATTTGGTGGAGGAACAGGCCCTGTGGCGGT 498
Qy 1392 ACGAATCGACTACTATCACTCCGGGCAAAATGGAATTTGCACCCGATGTTGGGCGCAGCA 1451
Db 499 ACTAATGCAACCAAGTATACCCCGGTATTTGGAAATATGTACCCGATGTTAGAGCGGTG 558
Qy 1452 GAAGAGTATTTCTAATGATGTTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACA 1511
Db 559 GATGAAATACCTAATTAATGTTGGGTATTTGGTAAAGGTTGTGTGTCAGTCAGCCCGAAGCA 618
Qy 1512 CTTGTAGAACAGTAGAGCGGCGCGATTTGGTTTAAATTTGCAATGAAGACTGGGCA 1571
Db 619 ATCCGCAACAAATAACAGCGGCTGCTATAGGTCTTAAATAACATGAAGACTGGGCGCA 678
Qy 1572 ACACCAAGTGGATCGATCACTGCTTGAGCGTGGCAGATGAATACGATGTCGAAGTTGT 1631
Db 679 ACGCAATGGCAATTCACATTTGCTTAATGTGCGGATGAATGAATGATGATACAGTGGCT 738
Qy 1632 ATCCACACCGATACAGTCAATGAGGCGAGTTATGTAGATGACACCTTAAATGCAATGAAC 1691
Db 739 ATTCACTCTGACACTTAAATGAAGTGGTTTTTATGAAGAGACAGTAAAGCCATTGCC 798
Qy 1692 GGGCGGCAATCCATGCTTACCAATTTGAGGAGGCTGAGGACACTACCTGATGTT 1751
Db 799 GGTGAGTGTATTCATATTCATACCAAGCGGAGTGGTGTGTCATGCCCTGATGTT 858
Qy 1752 ATCAACATGGCAGCGAGTCAATATTTCTCCCTCCACACCCCACTATTCCCTAT 1811
Db 859 ATCAAGTCGGTAGAGAGCCCAATATTTTCTGTCATCAACCAACCAAGATGCTTAT 918
Qy 1812 ACCAATTAATCGTTGCAAGACACTTAGACATGCTCATGATGCCACCACTAGACAAA 1871
Db 919 ACCAATTAATCGTTGAGAGCACTTGTATGTTGATGTTGCTGTCTCATCATCTCGATCCC 978
Qy 1872 CGATCCGCGAGGATTTACAAATTTCTCAAGCCGATTCGCCCCGCTCTATCGCGGT 1931
Db 979 TCTATTCTGAAGATGGGCAATTTGCTGAATCTGTTATTCGTCGCAAAACCAATGCTGCA 1038
Qy 1932 GAAGATGCTCCATGATATGGGTGATCGGATGACAACTCGGATTCGCAAGCAATG 1991
Db 1039 GAAGATATCTTACATGATATGGGGCAATTTCCGTTGATGTCGTCAGACTCAAGCCATG 1098
Qy 1992 GGGCGTCAGCGCAAGTATTTCTCGAACTTTGGCAGACTCGGCAAGTAAAGTAAAGAA 2051
Db 1099 GGAAGATCGGAGAGTTATCTTACGCACTTTGGCAGTGTGCATTAATGAATGCAA 1158
Qy 2052 TTTGGTAAGCTTCTGAAGATGGCAAGATAACGATAATTTCCGATTTAAGCGCTACATC 2111
Db 1159 CGAGGCAATTTAGCGGTGATAGCGCAGATAATGATAATAATCGTATTAAAGCTTATATC 1218
Qy 2112 TCCAATACATTAACCCCGCTTTGACCAACCGCGTGGAGGATATATCGGCTCTGTG 2171
Db 1219 GCTAATAACAGATTAATTCGCGCACTGGCAGATGGCATTTGCTTACGCGTGGATCAATA 1278
Qy 2172 GAAGGCAAGATCGCGACTTGTGTGTGGAACTCTGCTTTTGGCGTAAACCC 2231
Db 1279 GAAAAGGTAACCTTGGCGATATCGTGCTATGGGATCTGCTTTCTTTGGCGTAAACCG 1338
Qy 2232 AAAATCGTATCAAGGCGGTATGGTGTCTTCTGAAATGGGCGATTTCTAACCGCTCT 2291

Db 1339 GCATTATATATAAAGGTGGTATGTCGCTTATGCGCCAAATGGGGGATATTTATGCGGCT 1398
Qy 2292 GTGCCACTCCCAACCCGCTTTTATTTACCGCGAAATGTTTGGGCATCACGCGAAGCGAAA 2351
Db 1399 ATTCAACACCGCAACCGGTTCAITTTATCGTCCAATGTATGCTGTTTAGGAAAAGCCAAA 1458
Qy 2352 TTTGACACGAGCATCACTTTTGTGTTTCCAAAAGTCGCTATGAAAATGGCGTGAAAAGAAAG 2411
Db 1459 TATCAACAGTCGATGATCTTTATGTCAAAAAGCGGTTATTTAGCGCGGAGTGCAGAAAAA 1518
Qy 2412 CTGGCTTTAGAGCGCAAGTTCTACCGGTCAAAAACCTGCGTAACATCAACCAAGAAAGAC 2471
Db 1519 TAGGCTTTAAAAGCTTAATTTGTCGTGGAGGCTGCTGTCATATCAAAAAGCTTCG 1578
Qy 2472 TTCAAGTTCAACGACAAAACCGCAAAAATCACCGTCGATCCGAAAACCTTTCCGAGGCTTT 2531
Db 1579 ATGATCCACAATAACTATGTTCTCATATCGAATTAGATCCACAACTTACATTGTTAAA 1638
Qy 2532 GTAGATGCAAACTCTGCACCTCTAAACCCACCTGCGAAGTGGCTTAGCCCGAGCGCTAC 2591
Db 1639 GCGGATGTTGTACCACTGTTTGTGAGCCAGCGACTGAATTACCGATGGCTCAACGCTAT 1698
Qy 2592 ACTTTCTTCTA 2602
Db 1699 TTCTTATTTTA 1709

RESULT 13

US-08-967-513-1
; Sequence 1, Application US/08967513
; Patent No. 5783436
; GENERAL INFORMATION:
; APPLICANT: Robert P. Hausinger
; TITLE OF INVENTION: Mutant Urease and Method
; TITLE OF INVENTION: of Use For Determination
; TITLE OF INVENTION: of Urea
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,513
; FILING DATE: 11-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/687,645
; FILING DATE: July 26, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-309
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; TELEX: No. 5783436e
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2400
; TYPE: nucleotides
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION: cdna

;
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE:
; ORGANISM: Klebsiella aerogenes
; STRAIN: CG253
; INDIVIDUAL ISOLATE:
; CELL TYPE: N/A
; FEATURES:
; NAME/KEY: cDNA encoding mutant urease
; NAME/KEY: 'H219Q
; LOCATION: Modification at position 1312 to
; LOCATION: glutamine
; IDENTIFICATION METHOD: Sequencing
; OTHER INFORMATION:
;
US-08-967-513-1

Query Match 21.4%; Score 617.2; DB 2; Length 2400;
Best Local Similarity 60.3%; Pred. No. 2.7e-143;
Matches 1048; Conservative 0; Mismatches 683; Indels 6; Gaps 1;

QY 883 GTAAGGAGACTCCCATGAAATGAAAAAACAAGAAATATGTAAATACCTACCGACCCACCA 942
DB 645 GGAGGTAACCGATGAGTAATATTTACGCCAGGCCATATGCCGATATGTTGGGCCCCACCG 704
QY 943 AAGCGATAAAGTCGCTTAGGAGATACCGATCTTTGGCGAGAAGTAGAATGACTATA 1002
DB 705 TCGCGCAAGAGTGGCGCTGGCAGATACCGAGCTGTGGATCGAGGTGGAGACGATTTGA 764
QY 1003 CCACCTATGCGGAAGAACTTAAATTTGCGCGGGTAAACTATCCGTAGAGGTATGGTC 1062
DB 765 CCACCTACGGGGAAGAGTCAAAATTCGCGCGCGCAAGTATCCGACGGCATGGGCC 824
QY 1063 AGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCACTAAAGCCGATGATTCG 1122
DB 825 AGGACAGATGCTGGCGCGCAGCTGTGTGACCTGTGTCTACCAACCGGTGATCGTCG 884
QY 1123 ACTACCGGGATTTACAAAGCCGACATTTGGATTAAACGCGAAATCATCGCATTTG 1182
DB 885 ATCACTGGGGGATCGTTAAGGCCGATATCGCGGTGAAGAGCGCGGATCTTCGCCATCG 944
QY 1183 GCAAGCAGGAACAGGACATGCAAGATGGCGTAAGCCCTCATATGTTGTTGGGTGG 1242
DB 945 GCAAGCGCGCAACCCCGACATCCAGCCCAAGCTGACC-----ATCCCATCGCGCGCTG 998
QY 1243 GCACAGAGCACTAGCAGGGAAGGTATGATTATACCGCTGGGGGAATCGATTACACACA 1302
DB 999 CGACGAGAGTATCGCGCGCGAAGAAATTTGACCGCGCGGGATCGATACCCATA 1058
QY 1303 CCACATTCCTTTCTCCACAAATTCCTTACCGCTCTAGCCCAATGGCGTTACACCAATGT 1362
DB 1059 TTCCTGGATCTGTCCGACGAGCGGAAGGCGCTGTTCTGTGGGTGACCAACCATGG 1118
QY 1363 TTGAGCGGCGACAGTCTGTAGATGGAGAGATGCGATCTACTATCACTCCGGGCAAT 1422
DB 1119 TCGCGCGCGCACCGCGCGCGCGCGCCACCATGCCACCCATGCAACCCCGCGCGCGT 1178
QY 1423 GGAATTTGCAACCGCATGTTGCGCGCAGCAGAGAGTATTTATCAATGTGGCTTTTGG 1482
DB 1179 GGTATATCTCACGATGCTCGAGCGCGCCAGAGCTGCGCGGTCAATATCGCCCTGCTGG 1238
QY 1483 GCAAGCGCAATAGCTCTAGCAAAAAAACAATTTAGAACAAAGTAGAAGCGGCGCGATTG 1542
DB 1239 GCAAGGMAACGTTTCTCAGCCGATGTCCTGCGGAGCAGGTGGCGGAGCGGCTTATTG 1298
QY 1543 GTTTTAAATGCAATGAGACTGGGGGCAACAACAAGTGGATCGATCACTGCTTAGCG 1602
DB 1299 GCCTGAAGATCCAAGAGGACTGGGGGCGCCACCCCGCGCGCGATCGACTGTGCGTTAACCG 1358
QY 1603 TGGCAGATGATACGATGTCGAATTTGTATCCACACCGATACAGTCAATGAGCGAGTT 1662
DB 1359 TCGCCGATGAAATGGACATCCAGTTCGCCCTGTCAGCAGCAGCCCTGAATGAAATCGCGTT 1418
QY 1663 ATGTAGATGACACCCCTAAATGCAATGAACGGCGCGCCATCCATCCATACCATTTGAGG 1722

RESULT 14

US-08-687-645B-1

; Sequence 1, Application US/08687645B

; Patent No. 5846752

; GENERAL INFORMATION:

; APPLICANT: Robert P. Hausinger

; TITLE OF INVENTION: Mutant Urease and Method

; TITLE OF INVENTION: of Use For Determination

; TITLE OF INVENTION: of Urea

;

DB 1419 TTGTGAAGACACCTTCGCGGCATCGGGGCGCACCATCCACACCTTCATATCCGAAG 1478
QY 1723 GAGCGGGTGGAGGACACTCACCTGATTTATCCATATGGCAGCGAGCTCAATATTCTAC 1782
DB 1479 GGGCGGGGGGCGCATGCGCGGACATCATCACCGCTGCGGCCACCCGGAACATTTTCG 1538
QY 1783 CCTCTCCACGACCCCGCACTATTCCCTATACCATTAATATAGGTTGAGAACACTTAGACA 1842
DB 1539 CGTCGTCCACCAACCCCAACGCTGCGCTTACACCTCAACACCATCATGATGAACATCTCGATA 1598
QY 1843 TGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGGATTTACAATTTTCTCAAA 1902
DB 1599 TGCTGATGTTGCGCACCATCTGGACCCCGGAACATCGCGAGGACGTGGCCCTTTGCCGAGT 1658
QY 1903 GCGGTATCCGCCCGCGCTTATTCGGCGCTGAAGATGTGTCTCATGATATGGGTGTGATCG 1962
DB 1659 CGCGATTCGCGGGGAACCATCGCTGCGGAAGACGTGTGTCACGATCTCGCGCGCTTCT 1718
QY 1963 CGATGACAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGGAAGTGTATTCCTCGAACTT 2022
DB 1719 CGCTCACCTCTCCGATTCGCAAGGCAATGGGCGCGCTCGGGGAAGTGAATTTCTCGCACCT 1778
QY 2023 GGCAGACTGCGGATGAAGTAATAAAGAAATTTGGTAAGCTTCTGAAGTGGCAAGATA 2082
DB 1779 GGCAGGTGCGCGCATCGCATGAAGTGCAGCGCGAGCGCTGGCGGAGGACCGGGGATA 1838
QY 2083 ACGATAATTTTCGCATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCC 2142
DB 1839 ACGACACTTTCGCGTGAAGCGCTACATCGCCAAATACACCATCAACCCCGCGCTGACCC 1898
QY 2143 ACGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGATCGCCGACTTGGTGGTGT 2202
DB 1899 ACGGCATCGCACAGCAAGTCGATCCATTGAGGTGGTAAAGCTGGCTGACCTCGTGGTCT 1958
QY 2203 GGAATCTCGCTTTTTCGCGTAAACCCCAAAATCGTGAATCAAAAGCGGTATGGTGTCT 2262
DB 1959 GGTCAACAGCTTCTTTCGCGCGTGAACCCGCGCACCGTGAATCAAAAGCGCGCATGATCGCCA 2018
QY 2263 TCTCTGAATGGCGGATTTCAACGGCTGTGCGCCACTCCCAACCGGTTTATTACCGCG 2322
DB 2019 TCGCGCGATGGCGGATATCAATGCTCTTATTCGACCCCGCGAGCGGTGCACTACCGCC 2078
QY 2323 AAATGTTTGGGCATCAACGCGAAGCGAAATTTGACACCAAGCATCACTTTTGTTCCAAAG 2382
DB 2079 CGATGTTTGGCGCTGGCGAGCGCCCGCCATCACTGCGCGCTCACCTTCTCTGTCGAGG 2138
QY 2383 TCGCCTATGAAATGGCGTGAAGAAAGTGGGCTTAGAGCGCCAAAGTTCTTACCGGTCA 2442
DB 2139 CGCGCGCAGCAATGGCGTTGCGAGCGGCTGAACCTGCGCAGCGCATCGCGGTGTGA 2198
QY 2443 AAAACTGCGGTAAACATCAACCAAGAAAGACTTCAAGTTCAACGCAAAAAGCGCAAAATCA 2502
DB 2199 AAGTTCGCGTACCGGTGCAGAAAGCCGACATGGTGCACAAACAGTCTGCAAGCTTAAACATCA 2258
QY 2503 CCGTTCGATCCGAAACCTTTCGAGGTCTTTGATAGTGGCAAACTCTGCACTCTTAAACCCCA 2562
DB 2259 CCGTCGAGCCGCAACCTATGAGGTGCGGTGATGGCGAATTTATCAACGAGGCGCG 2318
QY 2563 CCTCGCAAGTCTCTAGCCGAGCGCTACATTTCTTCTTAGGCACAATGCCCCCTTT 2619
DB 2319 CAGACGTTCTGCCGATGGCGCAACGATATTTCTGTTTAAAGGAGAGCGGATGCTTT 2375

RESULT 14

US-08-687-645B-1

; Sequence 1, Application US/08687645B

; Patent No. 5846752

; GENERAL INFORMATION:

; APPLICANT: Robert P. Hausinger

; TITLE OF INVENTION: Mutant Urease and Method

; TITLE OF INVENTION: of Use For Determination

; TITLE OF INVENTION: of Urea

;

NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway
CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: wordperfect 5.1
CURRENT APPLICATION DATA:
FILING DATE: July 26, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MSU 4.1-309
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
TELEX: No. 5846752e
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2400
TYPE: nucleotides
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE:
DESCRIPTION: cdna
HYPOTHETICAL: No
ANTI-SENSE: No
ORIGINAL SOURCE:
ORGANISM: Klebsiella aerogenes
STRAIN: CG253
INDIVIDUAL ISOLATE:
CELL TYPE: N/A
FEATURE:
NAME/KEY: cdna encoding mutant urease
NAME/KEY: 'H219Q
LOCATION: Modification at position 1312 to
LOCATION: glutamine
IDENTIFICATION METHOD: Sequencing
OTHER INFORMATION:
US-08-687-645B-1
Query Match 21.4%; Score 617.2; DB 2; Length 2400;
Best Local Similarity 60.3%; Pred. No. 2.7e-143;
Matches 1048; Conservative 0; Mismatches 683; Indels 6; Gaps 1;
Qy 883 GTAAGGAGCTCCCATGAAATGAAAACAGATATGTAATACCTACGACCCACCA 942
Db 645 GGAGGTAACAGTAGTAAATTTTCACGCCAGGCTATGCCGATATGTTGGCCCCACCG 704
Qy 943 AAGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGAGTAGAACAATGACTATA 1002
Db 705 TCGGGACAGGTGGCCCTGCGAGATACCGAGCTGTGGATCGAGGTGGAGACGATTTGA 764
Qy 1003 CCACCTATGGCGAAGAACTTAAATTTGGCGGGTAAACCTATCCGTGAGGGTATGGTTC 1062
Db 765 CCACTACGGGAGAGGTCAAATTTGGCGGGCGCAAGTAGTCATCGCGACGGCATGGGCC 824
Qy 1063 AGAGCAATAGCCCTGATGAACACCCCTAGATTTAGTTCATCACTAACCGGATGATTTCG 1122
Db 825 AGGACAGATGCTGGCCCGCGACTGTGCGACCTGTGCTGCCAACCAGCGTTGATCGTCG 884

Qy 1123 ACTACACGGGATTTTCAAGCCGACATTGGGATTAATAAACCGCAAAATCCATGGCAATTG 1182
Db 885 ATCACTGGGGATCGTTAAGCCCGATATCGGCGTGAAGACCGCGCGGATCTTTCGCCATCG 944
Qy 1183 GCAAGGCAAGGAAACAAAGGACATGCAAGATGGGTAAGCCTCATATGGTCTGGTGGTGG 1242
Db 945 GCAAGCCGGCAACCCCGACATCCAGCCCAAGTCACTC-----ATCCCCATCGGCGCTG 998
Qy 1243 GCACAGAAAGCACTAGCAGGGGAAGTATGATTTATTAACCGTGGGGGAATTCGATTCACACA 1302
Db 999 CGACGGAAGTATCGCCCGCGAAGGAAATTTGTACCGCGCGGATCGATACCCATA 1058
Qy 1303 CCCACTTCTTCTCCCAACAAATTCCTACCGCTCTAGCCAAATGCGGTTCACCAATGT 1362
Db 1059 TTCCTGGATCTGTCCGAGCAGAGCGCTGTCTCTGCGCTGACCACTGG 1118
Qy 1363 TTGGAGGCGGCACAGGTCTCTGTAGATGGCAAGTCCGACTACTATCACTCCGGGCAAAAT 1422
Db 1119 TCGGCGGCGCACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1178
Qy 1423 GGAATCTGCACCGCATGTTGCGCGCAGCAGAAAGATTTCTATGAATGTGGGCTTTTGG 1482
Db 1179 GGTATATCTCAGCATGCTGCAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1238
Qy 1483 GCAAGGCAATAGCTCTAGCAAAAACAACTTGTAGAACAAAGTAGAGCGGGCGCGATTG 1542
Db 1239 GCAAGGGAACGTTTTCTCAGCGGATGCCCTCGCGAGCAGGTGGCGGCGGCGGCGGCTATT 1298
Qy 1543 GTTTTAAATTCATGAAGACTGGGCGCACAAACCAAGTCCGATCGATCACTGCTCGAGCG 1602
Db 1299 GCCTGAAGATCAAGAGGACTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1358
Qy 1603 TGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCGAGTT 1662
Db 1359 TCGCGATGAATGACATCCAGGTGCGCCCTGCACAGCGACACCCCTGAATGAATCGGTT 1418
Qy 1663 ATGTAGATGACACCCCTAAATGAAGACGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1722
Db 1419 TTGTGGAAGACACCCCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1478
Qy 1723 GAGCGGTGGAGGACACTCACCTGATGTTATCACCATGCGCGGCGGCGGCGGCGGCGGCGG 1782
Db 1479 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1538
Qy 1783 CCTCTCCACCAACCCCTACTATTCCCTATACCAATTAATACGGTTGCGAGAACCTTAGACA 1842
Db 1539 CGTGTCCACCAACCCCTAGCGTCCCTACACCTCAACACCATCGATGAACATCTCGATA 1598
Qy 1843 TGCTCATGACATGCCACCACTAGACAAACGGATCGCGGAGGATTTACAATTTCTCAAA 1902
Db 1599 TGCTGATGTTCTGCCACCATCTGGACCCGACATCGCCGAGGACGTTGGCGGCTTTTGG 1658
Qy 1903 GCGTATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1962
Db 1659 GCGCATTTGCGCGGGAACCATCGCTGCGGGAAGAGCGTGTGACAGATCTCGGCGGCTTCT 1718
Qy 1963 CGATCACAAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGCGGAAGTGATTTCTCGAATCT 2022
Db 1719 CGCTCACCTCTCCGATTCGAGGCCATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1778
Qy 2023 GCGAGCTCGGATGAAGATAAAGAAATTTGGTAAAGCTTCTGAAGATGGCAAGATA 2082
Db 1779 GCGAGTGGCGCATCGCATGAAGGTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGATA 1838
Qy 2083 AGGATAATTTCCGCAATTAAGCCCTACATCTCCAAATACACTATCAACCCCGCTTTGACCC 2142
Db 1839 ACACAACCTTCGCGGTGAAGCGCTACATCGCAAAATACACCATCAACCCCGGCGCTGACCC 1898
Qy 2143 ACGGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGGCAAGATCGCGGCGGCTTGGTGT 2202
Db 1899 ACGGCTGCGACACGAGTCCGATTCATTTAGGTGGGTAGCTGGCTGACCTCGTGTCT 1958
Qy 2203 GGAATCTCGCTTTTGGCGGTAACCAACCCAAATCGTGATCAAAAGCGGCGGATGGTGTCT 2262

Db 1959 GGTACACAGCTTCTTCGGCGTGAACCGGCCACCGTGATCAAGGGCGCATGTGCCA 2018
Qy 2263 TCTCTGAATGGGCGATTTCTAACCGCTCTGTGCGCACTCCCAACCGGTTTATTACCGCG 2322
Db 2019 TCGCGCGATGGGCGATATCAATGCCCTTATTCCGACCCCGCAGCGGTGCACTACCGCC 2078
Qy 2323 AATGTTTGGGCATCAGCGCAAGGCGAAATTTGACACAGCATCACTTTTGTTCGAAG 2382
Db 2079 CGATGTTTGGCGGCTGGGCGACGCCGCCCATCACTGCGGCTCACCTTCTGTGCGAGG 2138
Qy 2383 TCGCTATGAATGGCGTGAAAGAAAAGCTGGGCTTAGAGCGCAAGTCTCTACCGGTCA 2442
Db 2139 CGGGCGAGCAATGGGCTTGCAGCGGCTGAACCTTGGCGAGCGGATCGCGGTGTGA 2198
Qy 2443 AAAATGCCGTAAATCACCACCAAGAAAGACTTCAAGTTCAACGACAAAAACGCAAAAATCA 2502
Db 2199 AAGCTGCCGTACGGTGCAGAAAGCCGACATGGTGACACAGTCTGCAGCCTTAACATCA 2258
Qy 2503 CGGTGATCCGAAACCTTTCGAGTCTTTGTAGATGGCAAACTCTGCACCTCTTAAACCCA 2562
Db 2259 CCGTGCAGCGCCAGACCTATGAGGTGGGGTGGATGGCGAACTTATCACCAAGCGAGCGG 2318
Qy 2563 CTTGCGAAGTGCCTTAGCCAGCGCTACACTTTCTTCTAGGCAATGCCCCCTTT 2619
Db 2319 CAGAGCTTCTGCCGATGGCGAACGATATTTCTGTTTTAAGGAGCGGATGCTTT 2375

RESULT 15
US-09-489-039A-2045
; Sequence 2045, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2045
; LENGTH: 1878
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2045

Query Match 21.3%; Score 612.8; DB 3; Length 1878;
Best Local Similarity 60.8%; Pred. No. 3e-142;
Matches 1028; Conservative 0; Mismatches 657; Indels 6; Gaps 1;

Qy 900 AATGAAACCAAGATATGTAATACCTACGGACCCACCAAGCGGATAAAGTGGCG 959
Db 10 AATATTTCAGCGGCGCTATGCGGATATGTTGCGGCCACCGTGGCGCAAGGTGCGC 69
Qy 960 TTAGGAGATACCGATCTTTTGGCGAGAGTAGAATCACTATACCATCTATGGCGAAGAA 1019
Db 70 CTGGCGGATACCGAGCTGTGGATCGAGGTGAGGACGATCTGACCACTACGGGGAAGAG 129
Qy 1020 CTTTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGAT 1079
Db 130 GTCANATTCGGCGCGGCAAGTGATCCGCGACGGCATGGCGAGGGGAGATGCTGGCC 189
Qy 1080 GAAAAACCCCTAGATTAGTCATCACTAACCGGATGATTATCGACTACACGGGATTTAC 1139
Db 190 GCGCACTGTGTCACCTGTGTCTCACCAACGCGTTGATCGTCACTCACTGGGGGATCGTT 249
Qy 1140 AAGCGCATTTGGGATTAATAACGCGCAAAATCCATGGCATTTGCAAGGAGGCAAGCAAG 1199
Db 250 AAGGCGGATATCGGCTGGAAGGACGCGCGGATCTTGCCCATCGGCAAGCGGTAACCCC 309
Qy 1200 GACATGCAAGATGGCGTAAGCCCTCATATATGTTGTTGGTGTGGGACAGAACACTAGCA 1259

Db 310 GATATCCAGCCCAACGTCACC-----ATCCCATCGGTGCTCGACGGAAGTGTGCC 363
Qy 1260 GGGGAAGGTATGATTTATACCGCTGGGGGAATCGATTCACACACCACTTCTTCTTCCA 1319
Db 364 GCGGAAGGAATAATTTGTACCGCCGCGGATCGATACCATATTCACTGGATCTGTCCG 423
Qy 1320 CACAAATTTCCCTACCGCTCTAGCCATAGCGGTTACAAACATGTTTGGAGGCGCACAGGT 1379
Db 424 CAGCAGGCGGAAGAGGCGCTGTCTCCGGCGTGAACACCATGTTGCGGCGGACCGCGC 483
Qy 1380 CCGTGTAGATGGCAGAAATCGGACTACTATCACTCCGGGCAAAATGGAATTTGACACCGCATG 1439
Db 484 CCGGCGGCGGACCCATGCGCACCTGCAACCCCGGGCCCGTGTATATCTCAGCGATG 543
Qy 1440 TTGCGCGCAGCAGAAAGAGTATTCTATGATGTGGGCTTTTGGGCAAAAGCAATAGCTCT 1499
Db 544 CTGACGCGCGCGACAGCGCTCGGCTCAATATCGGCTGCTGGGCAAGGCAACGTTTCT 603
Qy 1500 AGCAAAAAAACAATTGTGTAGAACAGTAGAGCGGGCGCGATGGTTTAAATTTGCATGAA 1559
Db 604 CAGCGGATGCCCTCGCGGAGCAGGTGGCGGGCGGCTTATTGGCCCTGAAGATCCATGAG 663
Qy 1560 GACTGGGCGCAACACCAAGTGCAGTCACTGCTTCAGCGGTGCGAGATGAATAGAT 1619
Db 664 GACTGGGCGCGCCACCCCGGCGCGATCGACTGTGCGTTAACCGTCCCGGATGAATGGAC 723
Qy 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCAAGGTTATGTAGATGACACCTTA 1679
Db 724 ATCCAGTTCGCCCTGCACAGCGACACCTCGTATGATCGGTTTGTGGAGACACCTC 783
Qy 1680 AATCAATGAACGGCGCGCCATCATGCTACCACTTACCAATTTAGAGGAGCGGTGAGGACAC 1739
Db 784 GCGGCCATCGGCGGCGCACCATCCACACCTTCCATACCGAAGGGCGCGCGCGCCAT 843
Qy 1740 TCACCTGATTTATCACCATGCGAGCGAGCTCAATATTCTTACCTCTCCACCAACCC 1799
Db 844 GCGCGGACATCATCACCAGCTTCGCCCCACCCGAAACATTTTGGCGTCTCCACCAACCCA 903
Qy 1800 ACTATTCCTATACCATTAATACGGTTGCAGAACACTTAGACATGCTCATGATGATGCCAC 1859
Db 904 AGCTGCCCTTACACCTCAACACCATGATGAACATCTCGATATGCTGATGTGTGCCAC 963
Qy 1860 CACCTAGACAAAACGATCCGCGAGGATTTACAATTTTCTCAAAGCGGTATCCGCCCGCGC 1919
Db 964 CATCTGACCCGACATCGCGAGGACGTGGGCTTTTGCAGTTCGCGCATTCGCCGGAA 1023
Qy 1920 TCTATCGCGGTGAAGATGCTCATGATATGGTGTGATCGGATGACAGCTCGGAT 1979
Db 1024 ACCATCGCTGCGGAAGACGTGTGACAGATCTCGGTGCTTCTCGCTCACCTCCTCCGAT 1083
Qy 1980 TCGCAAGCAATGGGCGTGCAGGCGAAGTATTCCTCGAACTTTGGCAGATGCGGATAAG 2039
Db 1084 TCGCAGGCGATGGGCGCGTGGGGAAGTATTTCTCGCACCTTGGCAGGTGGCGCATCGC 1143
Qy 2040 AATAAAGAAGATTTGGTAAGCTTCTCAAGATGGCAAGATAACGATAATTTCCGCATT 2099
Db 1144 ATGAAGGTGACGCGGAGCGCTGGCGAGGAGACCGGGGATAACGACAACTTCCGGTG 1203
Qy 2100 AAGCGCTATCTTCCAAATACATTAACACCCGCTTTTGACCCACGCGGTGAGCGAGTAT 2159
Db 1204 AAGCGTACATCGCCAAATACCATCAACCGCGGCTGACCCACGCGCATCGCACAGAA 1263
Qy 2160 ATCGGCTGTGGAGAGGCGAAGATCCCGACTTGGTGTGTTGGAATCCTGCTTTT 2219
Db 1264 GTCGGATCCATTGAGGTGGGTAAAGCTGGTGAACCTCGTGGTGTGTCACGAGCTTCTTC 1323
Qy 2220 GCGTAAAAACCAAAATCGTGATCAAGCGGTATGGGTCTTCTCTGAAATGGGCGAT 2279
Db 1324 GCGGTGAACCGGCGCACCGTATCAAGCGGCGATGATCGCCATCGCGCGGATGGCGAT 1383
Qy 2280 TCTAACCGCTGTGTGCCCACTCCCAACCGGTTTATTATCCGCGAAATGTTTGGGATCAC 2339

```

Db      1384 ATCAATGCCTCTATTTCGACCCCGCAGCGGTGCACTACCGCCCGATGTTTGGCGCGTG 1443
Qy      2340 GGCAAGCGGAAATTTGACACCGACATCACTTTTTCCTCAAGTCGCTATGAAAATGGC 2399
      ||| ||
Db      1444 GGCAGCGCCCGCCATCACTGCGCCTCACCTTCCTGTGCGAGCGCGGCGCAATGGC 1503
      ||| ||
Qy      2400 GTGAAAGAAAAGCTGGGCTTAGAGCGGCAAGTTCTACGGTCAAAAACCTGCCGTAAACATC 2459
      || ||
Db      1504 GTTCCCGAGCGGCTGAACCTGCGCAGCGCGATCGCCGTGTGAAAGGCTGCCGTACGGTG 1563
      || ||
Qy      2460 ACCAAGAAAGACTTCAAGTTCAACGACAAAAACGGCAAAAATCACCGTCGATCCGAAACCC 2519
      || ||
Db      1564 CAGAAAGCCGACATGGTGCAACAGCCTGCGAGCTAACATCACCGTGCAGCGCCAGACC 1623
      || ||
Qy      2520 TTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAAGTCCTCTA 2579
      || ||
Db      1624 TATCAGGTGCGGTGGATGGCACTTATCACAGCGAGCGCGCAGAGGTTCTGCCGATG 1683
      || ||
Qy      2580 GCCAGCGCTA 2590
      || ||
Db      1684 GCGCAACGATA 1694
      || ||

```

Search completed: August 10, 2006, 02:19:22
 Job time : 542 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2006, 01:39:48 ; Search time 4662 Seconds

(without alignments)
7598.724 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2880.6

Sequence: 1 rrgagattttccarctt.....aaaaagtagaacacagg 2883

Scoring table: IDENTITY_NUC

Gapop 3.0 , Gapext 3.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main.*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10J_PUBCOMB.seq.*
- 16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10K_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2880.6	100.0	2883	US-09-904-994B-1	Sequence 1, Appli
2	2228	77.3	2452	US-09-904-994B-13	Sequence 13, Appli
3	2158.6	74.9	2405	US-09-904-994B-4	Sequence 4, Appli
4	2158.6	74.9	2407	US-09-904-994B-10	Sequence 10, Appli
5	1948	67.6	2183	US-09-904-994B-7	Sequence 7, Appli
6	1134.4	39.4	8407	US-10-639-273-1	Sequence 1, Appli
7	940.8	32.7	1710	US-10-476-313-1	Sequence 4, Appli
8	902.4	31.3	1719	US-10-476-313-10	Sequence 10, Appli
9	886.4	30.8	1710	US-10-335-977-3849	Sequence 3849, Ap
10	885.8	30.8	1815	US-09-995-913A-251	Sequence 251, App
11	884.8	30.7	1710	US-10-282-122A-22427	Sequence 22427, A
12	880	30.5	1710	US-10-500-447A-5	Sequence 5, Appli
13	876.8	30.4	1717	US-10-476-313-11	Sequence 11, Appli
14	860.6	29.9	2395	US-09-402-100-1	Sequence 1, Appli
15	665.4	23.1	1830121	US-10-329-670-1	Sequence 1, Appli
16	665.4	23.1	1830121	US-10-158-865-1	Sequence 1, Appli
17	665.4	23.1	1830121	US-10-981-687-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-904-994B-1
; Sequence 1, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1 N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodsequences
; CURRENT APPLICATION NUMBER: US/09/904, 994B
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2883
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (206)..(886)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (897)..(2603)
US-09-904-994B-1

Query Match 100.0%; Score 2880.6; DB 3; Length 2883;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RGGAGATTTTCCARCACTTCAAGCACATATTGATCTGTGTGGTGGTAAATTCR 60
DB 1 RGGAGATTTTCCARCACTTCAAGCACATATTGATCTGTGTGGTGGTAAATTCR 60
QY 61 ACTTGTAACTCTATTATATATTTTAAATTAATTTTATCTATTATATATATATATTA 120
DB 61 ACTTGTAACTCTATTATATATTTTAAATTAATTTTATCTATTATATATATATTA 120
QY 121 TTACTTATATATAAGTTAATAAAGTACGAAATTTAGCACTATATATCCATTCCTT 180
DB 121 TTACTTATATATAAGTTAATAAAGTACGAAATTTAGCACTATATATCCATTCCTT 180

QY 181 TAAATTTAACACAAAGAGTAATAGGTGAATCTCACACCCAAAGACGAAGAAAAGTTCTT 240
Db 181 TAAATTTTAAACAAAGAGTAATAGGTGAATCTCACACCCAAAGACGAAGAAAAGTTCTT 240
QY 241 GTTATATATTGCGGGCGAAGTGGCTAGAAAGCGCAAGACGAGGGCTTTAAAGCTCAACCA 300
Db 241 GTTATATATTGCGGGCGAAGTGGCTAGAAAGCGCAAGACGAGGGCTTTAAAGCTCAACCA 300
QY 301 ACCGAAAGCCATTGCTTACATTAGTGGCCCATPATTTATGGAAGCGGCGCGTGGAAAAA 360
Db 301 ACCGAAAGCCATTGCTTACATTAGTGGCCCATPATTTATGGAAGCGGCGCGTGGAAAAA 360
QY 361 AACCGTTGCCAGCTTATGGAAGAGTGATGCACTTTTGAAGAAAGATGAAGTAATGCC 420
Db 361 AACCGTTGCCAGCTTATGGAAGAGTGATGCACTTTTGAAGAAAGATGAAGTAATGCC 420
QY 421 CGGGTGGGTAAATATGTTCCCGATCTAGGTGTAGAAGCCACCTTTCCTGATGTAAGAA 480
Db 421 CGGGTGGGTAAATATGTTCCCGATCTAGGTGTAGAAGCCACCTTTCCTGATGTAAGAA 480
QY 481 ACTTGTAACCTGTGAATTTGGCCCATCGAACCAGATGAGCACTTCAAGCGGCGCAAGTGAA 540
Db 481 ACTTGTAACCTGTGAATTTGGCCCATCGAACCAGATGAGCACTTCAAGCGGCGCAAGTGAA 540
QY 541 ATTTGGTTGGATAAAGACATCGAGCTCAATGCGAGGCAAGAAAGTAACCGAACTTGAGGT 600
Db 541 ATTTGGTTGGATAAAGACATCGAGCTCAATGCGAGGCAAGAAAGTAACCGAACTTGAGGT 600
QY 601 TACTAAATGAAGGCGCTAAATCTTGCATGTGGGTAGCCATTCCACTTCTTTGAAGCTAA 660
Db 601 TACTAAATGAAGGCGCTAAATCTTGCATGTGGGTAGCCATTCCACTTCTTTGAAGCTAA 660
QY 661 CAAGGCACTAAAAATTCGATCGTGAAGAAAGCCTATGGCAAGCCCTAGATATTCCTCTGG 720
Db 661 CAAGGCACTAAAAATTCGATCGTGAAGAAAGCCTATGGCAAGCCCTAGATATTCCTCTGG 720
QY 721 CAACACGCTACGCAATGGGGCAGACAAAACCCGCAAGTGCACTTCTTTGGTGG 780
Db 721 CAACACGCTACGCAATGGGGCAGACAAAACCCGCAAGTGCACTTCTTTGGTGG 780
QY 781 CAGTAAAAAAGTGAATGGCATGAACGGGCTTGTGAATTAACATCGCGGATGAACGCCATAA 840
Db 781 CAGTAAAAAAGTGAATGGCATGAACGGGCTTGTGAATTAACATCGCGGATGAACGCCATAA 840
QY 841 ACATAAAGCGCTTCACAAGGCGAAATCTACGGATTTATCAAGTAAGGAGACTCCCATGA 900
Db 841 ACATAAAGCGCTTCACAAGGCGAAATCTACGGATTTATCAAGTAAGGAGACTCCCATGA 900
QY 901 AAATGAAAAAACAAGAAATATGTAATACCTACGGAACCCCAAGCGGCGATAAAGTGCGCT 960
Db 901 AAATGAAAAAACAAGAAATATGTAATACCTACGGAACCCCAAGCGGCGATAAAGTGCGCT 960
QY 961 TAGGAGATACCGATCTTTGGGCAGAGTAGAAACATGACTATACACCTATGGCGAAGAAC 1020
Db 961 TAGGAGATACCGATCTTTGGGCAGAGTAGAAACATGACTATACACCTATGGCGAAGAAC 1020
QY 1021 TTAATTTGGCGCGGTAAATCTACCTAGGGGTATGGGTGAGGCAATAGCCCTGATG 1080
Db 1021 TTAATTTGGCGCGGTAAATCTACCTAGGGGTATGGGTGAGGCAATAGCCCTGATG 1080
QY 1081 AAAACACCTAGATTTAGTCACTACACGCGATGATTTATCGACTACACCGGATTTACA 1140
Db 1081 AAAACACCTAGATTTAGTCACTACACGCGATGATTTATCGACTACACCGGATTTACA 1140
QY 1141 AAGCCGACATTTGGGATTTAAAAACGGCAAAATTCATGGCAATTTGGCAGGAGAAACAAGG 1200
Db 1141 AAGCCGACATTTGGGATTTAAAAACGGCAAAATTCATGGCAATTTGGCAGGAGAAACAAGG 1200
QY 1201 ACATGCAAGATGGGTAGCCCTCATATGGTCTGGGTGGGCAAGACACTAGCAG 1260
Db 1201 ACATGCAAGATGGGTAGCCCTCATATGGTCTGGGTGGGCAAGACACTAGCAG 1260
QY 1261 GGAAGGATGATATTATACCGTGGGGGAATCGATTTCACACCCACTTCTTCTCCAC 1320

Db 1261 GGAAGGATGATATTATACCGTGGGGGAATCGATTTCACACCCACTTCTTCTTCCAC 1320
QY 1321 AACAAATTCCTACCCCTCTAGCCAAATGGCGTTTACAAACCAATGTTTGGAGCGGCACAGGTC 1380
Db 1321 AACAAATTCCTACCCCTCTAGCCAAATGGCGTTTACAAACCAATGTTTGGAGCGGCACAGGTC 1380
QY 1381 CTGTAGATGGCAACGAATGGCACTATATCATCTCCGGGCAAAATGGAACTTGACACCGATGT 1440
Db 1381 CTGTAGATGGCAACGAATGGCACTATATCATCTCCGGGCAAAATGGAACTTGACACCGATGT 1440
QY 1441 TCGCGCAGCAGAGAGTATCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCTA 1500
Db 1441 TCGCGCAGCAGAGAGTATCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCTA 1500
QY 1501 GCAAAAAACAACCTTGTAGAAACGAAGCGGGCGGATTTGTTTAAATTCATGAAG 1560
Db 1501 GCAAAAAACAACCTTGTAGAAACGAAGCGGGCGGATTTGTTTAAATTCATGAAG 1560
QY 1561 ACTGGGCAACAACAAGTGGATCGATCACTGTGTGAGCGTGGCAGATGAATACGATG 1620
Db 1561 ACTGGGCAACAACAAGTGGATCGATCACTGTGTGAGCGTGGCAGATGAATACGATG 1620
QY 1621 TGCAAGTTTGTATCCACACCGGATCAGTCAATGAGCAGGTTATGTAGATGACACCTAA 1680
Db 1621 TGCAAGTTTGTATCCACACCGGATCAGTCAATGAGCAGGTTATGTAGATGACACCTAA 1680
QY 1681 ATGCAATGAACGGGCGCCATCCATCCATACCAATTCAGGGAGCGGTGGAGGACACT 1740
Db 1681 ATGCAATGAACGGGCGCCATCCATCCATACCAATTCAGGGAGCGGTGGAGGACACT 1740
QY 1741 CACTGTATTTATCAACATGGCAGCGAGCTCAATTAFTCTACCTCTCCACACCCCA 1800
Db 1741 CACTGTATTTATCAACATGGCAGCGAGCTCAATTAFTCTACCTCTCCACACCCCA 1800
QY 1801 CTATTTCCCTATACCAATTAATACGGTTGAGAGACACTTAGACATGCTCATGACATGCCACC 1860
Db 1801 CTATTTCCCTATACCAATTAATACGGTTGAGAGACACTTAGACATGCTCATGACATGCCACC 1860
QY 1861 ACCTAGACAAACGCAATCGCGAGGATTTACAATTTCTCAAAGCCGTATCCGCCCGGCT 1920
Db 1861 ACCTAGACAAACGCAATCGCGAGGATTTACAATTTCTCAAAGCCGTATCCGCCCGGCT 1920
QY 1921 CTATCGCGCTCAAGATGTCTCCATGATATGGGTGATCGCGATGACAAGCTCGGAT 1980
Db 1921 CTATCGCGCTCAAGATGTCTCCATGATATGGGTGATCGCGATGACAAGCTCGGAT 1980
QY 1981 CGCAAGCAATGGGGCGTGAGGCAAGTATCTCGAACTTGGCAGACTGCGGATGAAGA 2040
Db 1981 CGCAAGCAATGGGGCGTGAGGCAAGTATCTCGAACTTGGCAGACTGCGGATGAAGA 2040
QY 2041 ATAAAAAAGAAATTTGGTAACTTCTGAGAGTGGCAAGATACGATAATTTCCGCATTA 2100
Db 2041 ATAAAAAAGAAATTTGGTAACTTCTGAGAGTGGCAAGATACGATAATTTCCGCATTA 2100
QY 2101 AGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGGCGTGAGGAGTATA 2160
Db 2101 AGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGGCGTGAGGAGTATA 2160
QY 2161 TCGGCTCTGTGGAAGGCAAGATCGCGACTTGGTGTGGAAATCTTGCCTTTTTT 2220
Db 2161 TCGGCTCTGTGGAAGGCAAGATCGCGACTTGGTGTGGAAATCTTGCCTTTTTT 2220
QY 2221 GGTAAAAACCCAAATCGTGATCAAGGCGGTATGGTGTCTCTGAAATGGCGGAT 2280
Db 2221 GGTAAAAACCCAAATCGTGATCAAGGCGGTATGGTGTCTCTGAAATGGCGGAT 2280
QY 2281 CTAAAGCGCTCTGTGCGCACTCCCAACCGGTTTATTAACCGCGAAATGTTTGGGCATCACG 2340
Db 2281 CTAAAGCGCTCTGTGCGCACTCCCAACCGGTTTATTAACCGCGAAATGTTTGGGCATCACG 2340
QY 2341 GCAAGCGAAATTTGACACCGATCACTTTTGTTCCTCAAGTCGCTATGAAATGGCG 2400

Db 2341 GCAAGCGAAATTTGACACAGCATCACTTTTGTTCCTCAAGTCGCTATGAAATGCG 2400
Qy 2401 TGAAGAAAAGCTGGCTTAGAGCGCCAGTTCTACCGTCAAAAATCGCGGTAAATCA 2460
Db 2401 TGAAGAAAAGCTGGCTTAGAGCGCCAGTTCTACCGTCAAAAATCGCGGTAAATCA 2460
Qy 2461 CCAAGAAAGACTTCAAGTTCACAGCAAAACCGCAAAATCACCCTGATCGGAAACCT 2520
Db 2461 CCAAGAAAGACTTCAAGTTCACAGCAAAACCGCAAAATCACCCTGATCGGAAACCT 2520
Qy 2521 TCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAG 2580
Db 2521 TCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAG 2580
Qy 2581 CCAGCGCTACATTTCTTCTAGGCAAAATGCGCCCTTTGGGGGAGGTATTTAGGAA 2640
Db 2581 CCAGCGCTACATTTCTTCTAGGCAAAATGCGCCCTTTGGGGGAGGTATTTAGGAA 2640
Qy 2641 TCTTCATCAACAGCAGCTGCAATCGGTCTTGGGTGCGATCGTCTTAAACAAAC 2700
Db 2641 TCTTCATCAACAGCAGCTGCAATCGGTCTTGGGTGCGATCGTCTTAAACAAAC 2700
Qy 2701 TTTTCATCTTTAAGCAATCGCAATTTTAAATTAATTTAAATTTCTTAAATTAATTTAT 2760
Db 2701 TTTTCATCTTTAAGCAATCGCAATTTTAAATTAATTTAAATTTCTTAAATTAATTTAT 2760
Qy 2761 TATGCCCTCTCAATTTTAAAGGAGAAATATCGGTAGTCTTTGGTATTTGCTATGCGGG 2820
Db 2761 TATGCCCTCTCAATTTTAAAGGAGAAATATCGGTAGTCTTTGGTATTTGCTATGCGGG 2820
Qy 2821 TTGTTGGTCTGGCGCAAGAGGTATTTGAACCCATCGCTCAAAAAGTAGAAGCCAC 2880
Db 2821 TTGTTGGTCTGGCGCAAGAGGTATTTGAACCCATCGCTCAAAAAGTAGAAGCCAC 2880
Qy 2881 AGG 2883
Db 2881 AGG 2883

RESULT 2
US-09-904-994B-13
; Sequence 13, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Depoedesequities
; CURRENT APPLICATION NUMBER: US/09/904,994B
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2452
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(728)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (739)..(2445)
US-09-904-994B-13

Query Match 77.3%; Score 2228; DB 3; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 159 AGGACTATATCCATTCCTTTTAAATTTAAACAAAGGAGTAAATAGGTGAAACTCACAC 218
Db 1 AGGACTATATCCATTCCTTTTAAATTTAAACAAAGGAGTAAATAGGTGAAACTCACAC 60
Qy 219 CCAAGAGCAAGAAAAGTTCTTGTATATATGCGGGCGAAGTGGCTAGAAAGCGCAAG 278

Db 61 CCAAGAGCAAGAAAAGTTCTTTTATATATATCGGCGAAGTGGCTAGAAAGCGCAAG 120
Qy 279 CAGAGGCTTAAAGCTCAACCAACCGGAGCCATTGCTTACATTAAGTGCCTATATATGG 338
Db 121 CAGAGGCTTAAAGCTCAACCAACCGGAGCCATTGCTTACATTAAGTGCCTATATATGG 180
Qy 339 ACAGAGCGCGCTGGAAAACCGTTGCGCCAGCTTATGGAAGAGTGCATGCATTTT 398
Db 181 ACAGAGCGCGCTGGAAAACCGTTGCGGAACTTATGGAAGAGTGCATGCATTTT 240
Qy 399 TGAAGAAAAGATGAAGTAATGCGCGGGTGGGTAATATGCTTCCCGATCTAGGTGTAGAAG 458
Db 241 TGAAGAAAAGATGAAGTAATGCGCGGGTGGGTAATATGCTTCCCGATCTAGGTGTAGAAG 300
Qy 459 CCACCTTTCTGTATGGTACGAACTTTGTAACTGTGAACTTTGGCCCATCGAACCCAGATGAGC 518
Db 301 CCACCTTTCCCGATGGCACCAAACTCGTAACTGTGAACTTTGGCCCATCGAACCCAGATGAGC 360
Qy 519 ACTTCAAGCGGCGAAGTGAATTTGGTTGCGATTAAGACATCGAGCTCAATGCGAGCA 578
Db 361 ACTTCAAGCGGCGAAGTGAATTTGGTTGCGATTAAGACATCGAGCTCAATGCGAGCA 420
Qy 579 AAGAAGTAACCGAACTTGAAGTTACTTAATGAAGGCGCTAAATCTTGCATGTGGGTAGCC 638
Db 421 AAGAAGTTACCGAACTAGAGTTACTTAAGAAAGGACCTTAATCTTGCATGTGGGTAGCC 480
Qy 639 ATTTCCACTTCTTTGAAGCTAAACAGGCACTAAATTCGATCGTGAAGAAAGCTATGGCA 698
Db 481 ATTTCCACTTCTTTGAAGCTAAACAGGCACTTAATTCGATCGTGAAGAAAGCTATGGCA 540
Qy 699 AACGCTTAGATATTCCTCTGCGCAACAGCTACGATTTGGGGCAGGACAAACCCGCAAG 758
Db 541 AACGCTTAGATATTCCTCTGCGCAACAGCTACGATTTGGGGCAGGACAAACCCGCAAG 600
Qy 759 TGCAGTTGATTTCTTCTGCTGCGCACTAAAGTGAATTTGGCATGAACCGGCTTGTGAATA 818
Db 601 TGCAGTTAATCTTCTGCGGTAGTAAAGTGAATTTGGCATGAACCGGCTTGTGAATA 660
Qy 819 ACATCGGCGATGAACCGCAATAACATAAAGCGCTTGAAGCGGCAAAATCTCACGGATTTA 878
Db 661 ATATTGCGGACGACGCTTAAACCAAGCGCTAGACAAAGCAAAATCTCACGGATTTA 720
Qy 879 TCAAGTAAGGAGACTCCCATGAAATGAAACAAAGAAATATGTAAATACCTACGAGCC 938
Db 721 TCAAGTAAGGAGACTCCCATGAAATGAAACAAAGAAATATGTAAATACCTACGAGCC 780
Qy 939 ACCAAGCGGATTAAGTGGCTTAGGACATCCGATCTTTGGGAGAGTAGAAGCATGAC 998
Db 781 ACCAAGCGGATTAAGTGGCTTAGGACATCCGATCTTTGGGAGAGTAGAAGCATGAC 840
Qy 999 TATACCACTATGCGGAAGAACTTAAATTTGGCGGGTAAACCTATCCGTGAGGGTATG 1058
Db 841 TATACCACTATGCGGAAGAACTTAAATTTGGCGGGTAAACCTATCCGTGAGGGTATG 900
Qy 1059 GGTGAGCAATAGCCCTGTAGTAAACCAACCTTAGATTTAGTATCATCACTAACCGCATGATT 1118
Db 901 GGTGAGCAATAGCCCTGTAGTAAACCAACCTTAGATTTAGTATCATCACTAACCGCATGATT 960
Qy 1119 ATGACTACACCGGATTTTCAAGCCGACATTTGGGATTTAAACCGGCAAAATTCATGCG 1178
Db 961 ATGACTACACCGGATTTTCAAGCCGACATTTGGGATTTAAACCGGCAAAATTCATGCG 1020
Qy 1179 ATTGGCAAGCAGGAAACAGGACATGCAAGTGGCGTAAGCCCTCATATGTCGTGGGT 1238
Db 1021 ATTGGCAAGCAGGAAACAGGACATGCAAGTGGCGTAAGCCCTCATATGTCGTGGGT 1080
Qy 1239 GTGGGCAACAGGACATGAGGCGGAGGTATGATTATTACCGCTGGGGGAAATTCGATTCA 1298
Db 1081 GTGGGCAACAGGACATGAGGCGGAGGTATGATTATTACCGCTGGGGGAAATTCGATTCA 1140
Qy 1299 CACACCACTTCTTCTTCTTCCACAAATTCCTTACCGCTTAGCGCAATGGCGGTTACAC 1358
Db 1141 CACACCACTTCTTCTTCTTCCACAAATTCCTTACCGCTTAGCGCAATGGCGGTTACAC 1200

```
QY 1359 ATGTTTGGAGGCGGCGCAGAGTCTCTGTAGATGGCAAGAAATGCGACTACTATCACTCCGGGC 1418
Db 1201 ATGTTTGGCGGTGGCAGAGGCCCGGTAGATGGCAGAAATGCGACTACTCACTCCGGGC 1260
QY 1419 AATGGAACTTGCACCGCATGTTGGCGGCGAGCAAGAGTATTTCTATGAATGTGGGCTTT 1478
Db 1261 AATGGAACTTGCACCGCATGTTGGCGGCGAGCAAGAGTATTTCTATGAATGTGGGCTTT 1320
QY 1479 TTGGGCAAGGCAATAGCTCTAGCAAAAAAACAATTTAGAACAAAGTAGAAGCGGCGCG 1538
Db 1321 TTGGGCAAGGCAATAGCTCTAGTAAAAAACAATTTAGAACAAAGTAGAAGCGGCGCG 1380
QY 1539 ATTGGTTTTAAATGCAAGAGCTGGGGGCAACAACAAGTGCATGCATCTGCTTG 1598
Db 1381 ATTGGTTTTAAATGCAAGAGCTGGGGGCAACAATCCAAGTGCATGCATCTGCTTG 1440
QY 1599 AGCGTGCAGATGATGATGCGATGCGCAATGTTGTATCCACACGATACAGTCAATGAGGCA 1658
Db 1441 AGCGTAGCAGATGATGATGCGATGCGCAATGTTGTATCCACACGATACAGTCAATGAGGCA 1500
QY 1659 GGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGCGCATCCATGCGCTACCACTT 1718
Db 1501 GGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGCGCATCCATGCGCTACCACTT 1560
QY 1719 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACTATGCGAGGCGAGCTCAATATT 1778
Db 1561 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACTATGCGAGGCGAGTGAATATT 1620
QY 1779 CTACCCCTCTCCACCCACCCCACTATTCCCTATACCAATTAATAGGTTGCAAGACACTTA 1838
Db 1621 CTACCCCTCTCCACACCCCTACTATCCCTATACCAATTAATAGGTTGCAAGACACTTA 1680
QY 1839 GACATGCTCATGACATGCCACCTAGACAAACGATCCGCGAGGATTTTACAATTTCT 1898
Db 1681 GACATGCTTATGACTGCCACCTAGATAAAGCATCCGCGAGGATCTCCAATTTCT 1740
QY 1899 CAAAGCGGTATCCCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCATGATATGGGTG 1958
Db 1741 CAAAGCGGTATCCCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCATGATATCGGTG 1800
QY 1959 ATCCGATGACAAGCTCGATATCCGATGCAAGCAATGGGCGGTGACGGCAAGTGAATCTCGA 2018
Db 1801 ATCCGATGACAAGTTCGATTCGCAAGCAATGGGCGGCGCTGGGGAAGTGAATCTCTAGA 1860
QY 2019 ACTTGGCAGCTGGGATAGAAATAAAGAAATTTGGTAGACTTCTGAGAGTGGCAAA 2078
Db 1861 ACTTGGCAAACTGCAGACAAGAAATAAAGAAATTTGGTAGACTTCTGAGAGTGGTGA 1920
QY 2079 GATAACGATTAATTTCCGATTAAGCGCTACATCTCCAAATACATATCAACCCCGCTTG 2138
Db 1921 GATAATGACAACTTCCGATCAAGCGTATATCTCCAAATACACATTAATCCCGCTTG 1980
QY 2139 ACCACCGGTGACGAGTATATCGGCTCTGTGGAAGGCGCAAGATCCCGACTTGGT 2198
Db 1981 ACCATGGGTGACGAGTATATCGGCTCTGTGGAAGGCGCAAGATCCCGACTTGGT 2040
QY 2199 GTGTGGAATCTGCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAAGGCGGTATG 2258
Db 2041 GTGTGGAATCTGCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAAGGCGGTATG 2100
QY 2259 GTCTTCTGTAATGGCGATTTCTAAGCGCTGTGCGCCACTCCCCCAACCGGTTTATTAC 2318
Db 2101 GTGTTCTGTAATGGCGGATTTCTAATGCGTGTGTGCCACTCTCAGCCCGGTTATTAC 2160
QY 2319 CGCGAAATGTTTGGGCGATCACGGCAAGCGAAATTTTGACACAGCATCACTTTGTTTCC 2378
Db 2161 CGCGAAATGTTTGGGCGATCACGGCAAGCGAAATTTTGACACAGCATCACTTTGTTTCC 2220
QY 2379 AAGTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGCGGCAAGTTCACCG 2438
Db 2221 AAGTCGCTATGAAATGGTGTGAAGAAAACTTAGGTTTAGAGCGCAAGGTGCTCCC 2280
```

```
QY 2439 GTCAAAACTCCGCTAATCATCAAGAAAGACTTCAAGTTCAAACGACAAACGCGCAAAA 2498
Db 2281 GTGAAAAACTGCGCTAATCATCAAGAAAGACTTCAAGTTCAAACGACAAACGCGCAAAA 2340
QY 2499 ATCACCGTCTGATCCGAAAACCTTCCAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2558
Db 2341 ATCACCGTCTGATCCGAAAACCTTCCAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2400
QY 2559 CCCACTCTGCAAGTCCCTCTAGCCCAAGCGCTACACTTTCTTAGGCACAAT 2610
Db 2401 CCCACTCTGCAAGTCCCTCTAGCCCAAGCGCTACACTTTCTTAGGCACAAT 2452

RESULT 3
US-09-904-994B-4
; Sequence 4, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2405
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(681)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (692)..(2398)
; US-09-904-994B-4
```

Query Match 74.9%; Score 2158.6; DB 3; Length 2405;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

```
QY 206 GTGAACTCTACACCCAAAGAGCAAGAAAGTTCTTGTATATATATCGCGGCGAGTGGCT 265
Db 1 GTGAACTCTACACCCAAAGAGCAAGAAAGTTCTTGTATATATATCGCGGCGAGTGGCT 60
QY 266 AGAAGCGCAAGCAGAGAGGCTTAAAGCTCAACCAACCGAAGCCATTCTTACATTAGT 325
Db 61 AGAAGCGCAAGCAGAGAGGCTTAAAGCTCAATCAACCCGAAGCCATTGCTTACATTAGT 120
QY 326 GCCCATATTATGACCAAGCGCGCTGGAAAAAACCCTTGCCTTATGGAAGAG 385
Db 121 GCCCATATTATGACAGAGCGCGCTGGCAAAAAAACCCTTGTCTGNACTTATGGAAGAA 180
QY 386 TGATGCACTTTTTCAAAAAAGATGAAGTAATGCCGCGGTGGGTAAATATGGTTCCCGAT 445
Db 181 TGATGCACTTTTTCAAAAAAGATGAGGTGATGCCGCTGGGGAAATATGGTCCCTGAT 240
QY 446 CTAGGTGTAGAGCCACCTTTCTGATGCTAGCAAACTTGTAACTGTGAATTCGCCCATC 505
Db 241 TTGGCGGTAGAAGCCACTTTCCCGATGGCACTCGTAACCGGTGAATTCGCCCAT 300
QY 506 GAAACAGATGACACTTCAAGCGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAG 565
Db 301 GAACTGATGAACACTTTTAAAGCGGTGAAGTGAATTTGGCTGTGATTAAGACATTTGAG 360
QY 566 CTCAATGACGCAAGAAAGTAACCGAAGTGTAGGTTTACTTAATGAAGGGCTTAATCTTG 625
Db 361 CTCAAGCGGCTAAGGAAGTTACCGAGCTTGAAGTTTACCAACGAAAGCACTTAATCTTG 420
QY 626 CATGTGGGTAGCCATTTCCACTTTCTTGAAGCTTAACAGGCACATAAATTCGATCGTGA 685
Db 421 CATGTGGGTAGCCATTTCCACTTTCTTGAAGCAACCAAGGCATTTGAATTCGATCGGAA 480
```



```
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(682)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (693)..(2399)
US-09-904-994B-10.

Query Match      74.9%; Score 2158.6; DB 3; Length 2407;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 206 GTGAAAATCTCACCCAAAGAGCAAGAAAGTCTTTGTTATATATGCGGGCGAAGTGGCT 265
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2  GTGAAATCTCACCCAAAGAGCAAGAAAGTCTTTGTTATATATGCGGGCGAAGTGGCT 61
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 266 AGAAAGCGCAAGAGAGGGCTTAAAGCTCAACCAACCCGAAGCATTTGCTTACATAGT 325
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 62  AGAAAGCGCAAGAGAGGGCTTAAAGCTCAACCAACCCGAAGCATTTGCTTACATAGT 121
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 326 GCCATATTATGAGAGAGCGCGCGTGGAAAAAACCCTTGCCAGCTTATGGAAGAG 385
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 122 GCCATATTATGAGAGAGCGCGCGTGGAAAAAACCCTTGCCAGCTTATGGAAGAG 181
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 386 TGCATGCACTTTTGA AAAAAGATGAAGTAATGCCGGGTGGGTAAATATGTTCCCGAT 445
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 182 TGTATGCACTTTTGA AAAAAGAGAGGTGATGCCGGTGGGGAATATGTTCCCTGAT 241
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 446 CTAGGTGTAGAGCCACTTTCTTGATGTTACGAAACTTGAACCTGTAATGTCGCCATC 505
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 242 TTAGCGGTGGAAGCTACTTTTCCCGATGGCACCAAACTCGTAAACGTTGAAATTGGCCCATC 301
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 506 GAACCATATGAGCACTTCAAGCGGGCGAAGTGAATTTGTTGCGATAAAGACATCGAG 565
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 302 GAACCCGATGAACACTTCAAGCGGGCGAAGTCAAAATTTGGCTGTGTATGAAGACATTGAA 361
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 566 CTCATCGAGCAAGAGTAACCGAATCTCAGAGTTACTAATGAAGGCCCTAAATCCCTTG 625
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 362 CTCAGCGAGTAAGAGTTACCGAATAGAACTAGAACTTACCAAGAGGACCTTAAATCCCTTG 421
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 626 CATGTGGGTAGCCATTTCACCTTTCTTGAAGCTAACAGGCACTAAAAATTCGATCGTGA 685
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 422 CATGTGGGTAGCCATTTCACCTTTCTTGAAGCAACAGGCACTTGAATTCGATCGGGA 481
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 686 AAAGCCTATGCAAAACGGCTTAGATATTCCTCTGCGCAACACGCTACGCAATTTGGGGCAGGA 745
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 482 AAAGCCTATGCAAAACGGCTTAGATATTCCTCTGCGCAACACGCTACGCAATTTGGGGCAGGA 541
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 746 CAACCCGTAAGTGCAAGTTCCTCTTGGTGGCAGTAAAGTCAATTTGGCATGAAC 805
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 542 CAACCCGTAAGTGCAAGTTCCTCTTGGTGGCAGTAAAGTCAATTTGGCATGAAC 601
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 806 GGGCTTGTGAATAACATCGCGGATGAACCCATAAACATAAAGCGCTTTGA CAAGGCGAAA 865
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 602 GGGCTTGTGATATATTGCAAGTGAACGCCATAACACAAAGCGTTAGAAAAACCAAA 661
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 866 TCTCAGCGATTTCAAGTAAGAGAGACTCCCATGAAAAATGAAAAACAAGAAATATGTAAA 925
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 662 TCTCAGCGATTTCAAAATAAGAGAGACTCCCATGAAAAATGAAAAACAAGAGTATGTAAA 721
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 926 TACCTACGGACCCCAAGCGGATAAGTGGCTTAGAGATACCGATCTTTGGGCAGA 985
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 722 TACCTACGGACCTTACCAAGCGGCAAAAGTGGCTTAGAGATACCGATCTTTGGGCAGA 781
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 986 AGTAGAATCACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGGTAAAACTAT 1045
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 782 AGTAGAATCACTATACCACTATGGCGAAGAGCTCAAAATTTGGCGGGGTAAAACTAT 841
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1046 CCGTGAGGGTATGGGTCAAGACCAATAGCCCTGATGAAAAACCCCTPAGATTTAGTCATCAC 1105
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 842 CCGTGAGGGCATGGGTCAAGCAATAGTCCAGATGAAAACACCCTAGATTTAGTCATCAC 901
Qy 1106 TAACCGCATGATTATCGACTACACCGGGATTTTCAAAAGCCGCACTTTGGGATTTAAAAACGG 1165
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 902 CAACCGCATGATTATGACTACACCGGGATTTTCAAAAGCCGCACTTTGGGATTTAAAAATGG 961
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1166 CAAAATCCATGGCATTTGGCAAGGAGGAAAACAAGACATGCAAGATGGGCTAAGCCCTCA 1225
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 962 CAAAATCCATGGCATTTGGCAAGGAGGAAAACAAGACATGCAAGATGGGCTAAGCCCTCA 1021
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1236 TATGTCGTGGTGTGGGCACAGAAAGCACTAGCAGGGGAAGGTATGATTATTACCGCTGG 1285
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1022 TATGTCGTGGTGTGGGCACAGAAAGCATTAGCAGGGGAAGGTATGATTATTACCGCTGG 1081
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1286 GGGAAATCGATTACACACCCACTTCTCTTCTCCACAACAATTTCCCTACCGCTTAGCCAA 1345
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1082 GGGGATCGATTACACACCCACTTCTCTCTCTCCACAACAATTTCCCTACCGCTTAGCCAA 1141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1346 TGGCGTTACACCATTTTGGAGGCGGCACAGTCTGTAGATGGCACGAATGCGACTAC 1405
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1142 TGGCGTTACACCATTTTGGCGGTGGCACAGTCCGCTAGATGGCACGAATGCGACTAC 1201
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1406 TATCACTCCGGCAAAATGGAACCTTGCACCGCATTTGGCGCAGCAGAGAAGATTTCTAT 1465
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1202 CATCACTCCGGCAAAATGGAACCTTGCACCGCATTTGGCGCAGCAGAGAAGATTTCTAT 1261
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1466 GAATGTGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACCACTTTGTGAACAAGT 1525
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1262 GAATGTGGCTTTTGGGCAAAAGGCAATAGCTCCAGTAAAAAACCACTTTGTGAACAAGT 1321
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1526 AGAAGCGGGCGGATTTGTTTAAATTTGCATGAAGACTGGGGCAACAACAAGTGGGAT 1585
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1322 AGAAGCGGGCGGATTTGTTTAAATTTGCATGAAGACTGGGGCAACAACAAGTGGCAAT 1381
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1586 CGATCACTGCTTTGAGCGTGGCAGATGAATACCGATGTGCAAGTTTGTATCCACACCGATAC 1645
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1382 CGATCACTGCTTTGAGCGTAGCAGATGAATACCGATGTGCAAGTTTGTATCCACACCGATAC 1441
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1646 AGTCAATGAGGCAAGTTATGTAGATGACACACCCCTAAATGCAATGAACGGGCGGCCATCCA 1705
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1442 GGTCAATGAGGCAAGTTATGTAGATGACACACCCCTGAATGCGATGAACGGGCGGCCATCCA 1501
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1706 TGCCTAACCAATTTGAGGAGCGGGTGGAGGACACTCACTGATTTATCAACATGGCAGG 1765
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1502 TGCCTAACCAATTTGAGGAGCGGGTGGAGGACACTCACTGATTTATCAACATGGCAGG 1561
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1766 CGAGCTCAATATTTCTACCTCTCCACACCCCACTATTCCCTATACCATTAATACGCT 1825
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1562 CGAGCTCAATATTTCTACCTCTCCACCAACCCCACTATCCCTCTATACCATTAATACGCT 1621
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1826 TGCAGAACACTTTAGACATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGGA 1885
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1622 TGCAGAACACTTTAGACATGCTCATGACCTGCCACCACTAGATAAACGCAATCCGCGAGGA 1681
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1886 TTTTCAATTTTCTCAAAAGCGGTATCCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCA 1945
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1682 TTTTCAATTTTCTCAAAAGCGGTATCCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCA 1741
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1946 TCAATATGGTGTGATCGGATGACAGCTCGGATTCGAGCAATGCGGCGTGCAGCGGA 2005
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1742 TGATATTTGGCGTGTGATCGGATGACAGCTCGGATTCGAGCAATGCGGCGCTGCGCGA 1801
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2006 AGTGATTTCTCGAACTTTGGCAGACTGCGGATTAAGATAAAAAAGAAATTTGGTAAAGCTTCC 2065
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1802 AGTGATTTCTCGAACTTTGGCAGACTGCGGATTAAGATAAAAAAGAAATTTGGTAAAGCTTCC 1861
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2066 TGAAGATGSCAAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTAT 2125
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1862 TGAAGATAGTGAGATGAAGCAACTTCCGTATCAAAACGCTACATCTCCAAATACACTAT 1921
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2126 CAAACCCCGCTTTGACCCACGGGCTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAGAT 2185
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1922 TAAACCCCGCTCTAAACCCATGGGGTGAAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAAAT 1981
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


QY 2186 CGCGACTTGGTGTGGATCCTGCTTTTGGCGTAAACCCAAATCGTGATCAA 2245
DB |||||
QY 1982 CGCTGATTGGTGGTGGAAATCCTGCTTTTGGTGTGAACCTAAGATTGTGATCAA 2041
DB |||||
QY 2246 AGCGGGTATGGTGGTCTCTGAAATGGGGGATCTAAACGGCTGTGGCCCACTCCCA 2305
DB |||||
QY 2042 AGCGGGTATGGTGGTCTCTGAAATGGGGGATCTAAACGGCTGTGGCCCACTCA 2101
DB |||||
QY 2306 ACCGGTTATTACCGCGAAATGTTTGGGCATCAGCGCAAGCGAAATTTGACACCAT 2365
DB |||||
QY 2102 GCGGTTTATTACCGCGAAATGTTTGGGCATCAGCGCAAGCGAAATTTGACACCAT 2161
DB |||||
QY 2366 CACTTTTGTTCGAAGTCGCTATGAAATGGGGTGAAGAAAGCTGGGGCTTAGAGCG 2425
DB |||||
QY 2162 CACTTTTGTTCGAAGTCGCTATGAAATGGGGTGAAGAAAGCTTAGAGCG 2221
DB |||||
QY 2426 CCAAGTTCTACCGGTCAAAATCTGCGGTAAACATCAACCAAGAAAGCTTCAAGTTCAACGA 2485
DB |||||
QY 2222 CAAGGTGCTACCGGTGAATACTGCGCAACATCACTAAGAAAGCTTCAAAATCAACAA 2281
DB |||||
QY 2486 CAAACCGCAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGCGAACT 2545
DB |||||
QY 2282 CAAGCGCGCATATCACTGTCGATCTTAAACCTTCGAGGTCTTTGTAGATGCGAACT 2341
DB |||||
QY 2546 CTGCACCTCTAAACCCCACTCGCAAGTGCCTTAGCCAGCGCTACACTTTCTTAGGC 2605
DB |||||
QY 2342 CTGCACCTCTAAACCCCGCTCTGAAGTGCCTTAGCCAGCGCTACACTTTCTTAGGC 2401
DB |||||
QY 2606 ACAATG 2611
DB |||||
QY 2402 NCAATG 2407
DB |||||

RESULT 5

US-09-904-994B-7
; Sequence 7, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degosedesquences
; CURRENT APPLICATION NUMBER: US/09/904, 994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2183
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(683)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (694)..(2181)
US-09-904-994B-7

Query Match 67.6%; Score 1948; DB 3; Length 2183;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 206 GTGAACTCACCCCAAGAGCGAAGAAAGTCTTGTATATATATGCGGGAAGTGGCT 265
DB |||||
QY 3 GTGAACTCACCCCAAGAGCGAAGAAAGTCTTGTATATATGCGGGAAGTGGCT 62
DB |||||
QY 266 AGAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCCAAGCCATTTGCTTACATTA 325
DB |||||
QY 63 AGAAGCGCAAGCAGAGGGCTTAAAGCTCAATCAACCCCAAGCCATTTGCTTACATTA 122
DB |||||
QY 326 GCCATATTATGACGAGCGCGCTGGAAAAAACCCTTGTCCCAAGCTTATGGAAGAG 385
DB |||||
QY 123 GCCATATTATGACGAGCGCGCTGGAAAAAACCCTTGTGTAACCTTATGGAAGAA 182
DB |||||

QY 386 TGCATGCACCTTTTGAATAAGATGAATATGCCCGGGTGGGTAAATATGTTCCCGAT 445
DB |||||
QY 183 TGTATGCACCTTTTGAATAAGATGAATATGCCCGGGTGGGTAAATATGTTCCCGAT 242
DB |||||
QY 446 CTAGGTGTAGAAGCCACCTTTCTGATGTGTACCAAACTTCTAATCTGTGAATTTGGCCCATC 505
DB |||||
QY 243 TTGGGGGTAGAAGCCACCTTTCCCGATGGCAAACTCGTAACCGTGAATTTGGCCCAT 302
DB |||||
QY 506 GAACAGATGAGCACCTTCAAAGCGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAG 565
DB |||||
QY 303 GAACCTGATGAACACCTTTAAAGCGGCTGAAGTGAATTTGGCTGTGATAAAGACATTTGAG 362
DB |||||
QY 566 CTCAATGCAAGGCAAGAAGTAAACCGAATTTAGGTTACTTAATGAAGGCGCTAAATCTTTG 625
DB |||||
QY 363 CTCAACGCTGGGTGAAGGAAGTTTACCGAGCTTGAAGTTTACCAACGAAGGACCTAAATCTTTG 422
DB |||||
QY 626 CATGTGGGTGAGCCATTTCCACCTTTCTTGAAGCTTAAACAGGCACCTAAATTTGATCGTGAA 685
DB |||||
QY 423 CATGTGGGTGAGCCATTTCCACCTTTCTTGAACCAACAGGCATTTGAATTTGATCGGAA 482
DB |||||
QY 686 AAAGCTTATGCAAAACGCTTAGATATTCCTCTCTGGCAACACGCTAGCGATTTGGGGCAGGA 745
DB |||||
QY 483 AAAGCTTATGCAAAACGCTTAGATATTCCTCTCTGGCAACACGCTAGCGATTTGGGGCAGGA 542
DB |||||
QY 746 CAAACCCGCAAAAGTGCAGTTGATTTCTTGTGTGGCAGTAAATAAGTGTGATCGATGAAC 805
DB |||||
QY 543 CAAACCCGTTAAGTGCAGTTAATTCCTCTTGGCGGTAGTAAATAAGTGTGATCGATGAAC 602
DB |||||
QY 806 GGGCTTGTGAATAACATCGCGGATGAACCGCATTAACATTAAGCGCTTGAACAGGCGAAA 865
DB |||||
QY 603 GGGCTTGTGAATAATTTTGGCGCAACCGCATTAACATTAAGCGCTTGAACAGGCGAAA 662
DB |||||
QY 866 TCTCACGATTTTATCAAGTAAAGGAGACTCCCATGAATAATGAATAAACAAGATATGTAAA 925
DB |||||
QY 663 TCTCACGATTTTATCAAGTAAAGGAGACTCCCATGAATAATGAATAAACAAGATATGTAAA 722
DB |||||
QY 926 TACTTACGAGCCCAACAAAGCGGATAAAGTGGCTTTAGGAGATACCGATCTTTTGGGCAGA 985
DB |||||
QY 723 CACTTACGAGCCCAACCAAGCGGATAAAGTGGCTTTAGGAGATACCGATCTTTTGGGCAGA 782
DB |||||
QY 986 AGTAGAATGATGATATACCATTTAGCGGAAGAACTTAAATTTGGCGGGGTAAATCTAT 1045
DB |||||
QY 783 AGTAGAATGATGATATACCATTTAGCGGAAGAGCTCAAAATTTGGCGGGGTAAATCTAT 842
DB |||||
QY 1046 CGGTGAGGTATGGTCAAGCAATAGCCCTCATGAACACCCCTAGATTTAGTCATCAC 1105
DB |||||
QY 843 CCGTGAAGGTATGGGTCAAGCAATAGCCCAATAGCCCAATGAATAACCTTAGATTTAGTGATCAC 902
DB |||||
QY 1106 TAACCGATGATTTATCGACTACACCGGATTTTCAAGCGGACATTTGGGATTTAAATAACGG 1165
DB |||||
QY 903 CAACCGGATGATTTATCGACTACACCGGATTTTAAAGCGGACATTTGGTATTTAAATAAG 962
DB |||||
QY 1166 CAAATTCATCGCATTTGCAAGGCGAGGAAACAAAGGACATGCAAGATGGCGTGAAGCCCTCA 1225
DB |||||
QY 963 CAAATTCATCGCATTTGCAAGGCGGGAACAAAGACATGCAAGATGGCGTGAAGCCCTCA 1022
DB |||||
QY 1226 TATGTCGTGGTGTGGGCACAGAGCACTAGCAGGGGAGGTATGATTTATACCGCTCG 1285
DB |||||
QY 1023 TATGTCGTGGTGTGGGCACAGAGCACTAGCAGGGGAGGTATGATTTATACCGCTCG 1082
DB |||||
QY 1286 GGGATCGATTTACACACCCCTTCTTCTCCCAAACTTCCCTACCGCTCTAGCCAA 1345
DB |||||
QY 1083 GGGATCGATTTGCAACCCCTTCTTCTCCCAAACTTCCCTACCGCTCTAGCCAA 1142
DB |||||
QY 1346 TGGGCTTACAAACCATTTTGGAGGCGGACAGGTCTCTGTAGATGGCAGCAAGATGCTATAC 1405
DB |||||
QY 1143 TGGGCTTACAAACCATTTTGGAGGCGGACAGGTCTCGGTAGATGGCAGCAAGATGCTATAC 1202
DB |||||
QY 1406 TATCACTCCGGGCAAAATGGAATCTTGCAACCGCATTTGGCGCAGCAGAGATTTCTAT 1465
DB |||||
QY 1203 CATCACTCCGGGCAAAATGGAATCTTGCAACCGCATTTGGCGCAGCAGATTTCTAT 1262
DB |||||

1466 GAATGTGGCTTTTGGCAAGCAATAGCTAGCAAAAAACAACCTTTAGAACAGT 1525
 Db GAATGTAGCTTTTGGCAAGCAATAGCTTAGCAAAAAACAACCTTTAGAACAGT 1322
 QY AGAAGCGGGCGGATTTGTTTAAATTGCAAGAGCTGGGGCAACAACCAAGTGGAT 1585
 Db AGAAGCGGGCGGATTTGTTTAAATTGCAAGAGCTGGGGCAACAACCAAGTGGAT 1382
 QY CGATCACTGCTTGGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1645
 Db CGATCACTGCTTGGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1442
 QY AGTCAATGAGCAGGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1705
 Db GGTCAATGAGCAGGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1502
 QY TGCTTACCACTTGGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1765
 Db TGCTTACCACTTGGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1562
 QY CGAGCTCAATATTTACCTCTCCACCCACCCCACTATTCCTTATACCAATTAATACGGT 1825
 Db CGAGCTCAATATTTACCTCTCCACCCACCCCACTATTCCTTATACCAATTAATACGGT 1622
 QY TGCAAGACATTTAGACATGCTATGATGATGATGATGATGATGATGATGATGATGAT 1885
 Db TGCAAGACATTTAGACATGCTATGATGATGATGATGATGATGATGATGATGATGAT 1682
 QY TTTACAAATTTTCTCAAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1945
 Db TTTACAAATTTTCTCAAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1742
 QY TGATATGGGTGATCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2005
 Db TGATATGGGTGATCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1802
 QY AGTGAATCTCGAATTTGGGAGATGATGATGATGATGATGATGATGATGATGATGAT 2065
 Db AGTGAATCTCGAATTTGGGAGATGATGATGATGATGATGATGATGATGATGATGAT 1862
 QY TGAAGATGCGCAATTAACGATGATGATGATGATGATGATGATGATGATGATGAT 2125
 Db TGAAGATGCGCAATTAACGATGATGATGATGATGATGATGATGATGATGATGAT 1922
 QY CAACCGGCTTTGACCAATTAACGATGATGATGATGATGATGATGATGATGATGAT 2185
 Db TAATCCCGCTTTGACCAATTAACGATGATGATGATGATGATGATGATGATGATGAT 1982
 QY CGCGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2245
 Db CGCGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2042
 QY AGCGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2305
 Db AGCGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2102
 QY ACCGCTTTTATACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2365
 Db GCCGCTTTTATACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2162
 QY CACTTTTGTTCCTCAAGTCG 2385
 Db CACTTTTGTTCCTCAAGTCG 2182

RESULT 6
 US-10-639-273-1
 ; Sequence 1, Application US/10639273
 ; Publication No. US200401423431
 ; GENERAL INFORMATION:
 ; APPLICANT: Chang, Yung-Fu
 ; APPLICANT: Simpson, Kenneth W
 ; APPLICANT: Zhu, Jiaqian

; TITLE OF INVENTION: HELICOBACTER BIZZOZONII UREASE GENES AND THEIR USES
 ; FILE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
 ; FILE REFERENCE: 19603/3881
 ; CURRENT APPLICATION NUMBER: US/10/639,273
 ; PRIOR FILING DATE: 2003-08-12
 ; PRIOR FILING DATE: 2002-08-16
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 8407
 ; TYPE: DNA
 ; ORGANISM: Helicobacter bizzoeronii
 ; US-10-639-273-1

Query Match 39.4%; Score 1134.4; DB 8; Length 8407;
 Best Local Similarity 68.5%; Pred. No. 1.7e-238;
 Matches 1681; Conservative 0; Mismatches 741; Indels 32; Gaps 2;

QY 183 AAATTTAACACAGGAGTAATAGGTGAATACTCACCCCAAAGAGCAAGAAATTTCTGT 242
 Db 2319 AAATTTGGTAGAAGGAGTTTAGGATCAAAATTAAACCTTAAGAGCTGGACAAGCTCATGT 2378
 QY 243 TATATTATCGGGCGAAGTGGCTAGAAAGCGCAAGACAGAGGCTTAAAGCTCAACCAAC 302
 Db 2379 TGCATTATCGGGCGCAATTTGGCTTAAAAACGCAAGCAAAATGGCGTTAAGCTTAAATATA 2438
 QY 303 CGAAGCCATTTCCTTACATTTAGTGCCTTATTTATGACGAAGCGCGCTTGAAGAAAAA 362
 Db 2439 CTGAGCAGTAGCCCTCATCTAGTGCCTTATGATGGAAGAGCCCGTGCAGGTAAAAA 2498
 QY 363 CGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGTGAAAAAAGATGAATGATGCGCC 422
 Db 2499 GTGTGCGGATTTGATGCAAGAGGCGAGGACACTTCTTAAAGCTGATGATGTCTATGCGCG 2558
 QY 423 GGTGCGGTAAATGCTTCCGATCTAGGTGATGAGAGCCACTTCTCTGATGATGATGAGAAC 482
 Db 2559 GTGTAGCCCATATGATGCCAAGTGGGATTAAGAGCTAACTTCCCTGATGGACAAAA 2618
 QY 483 TTGTAACTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAGCGGCGCAAGTGAAT 542
 Db 2619 TGTTAAACCATCATACCCCGCTTGAAGATGGTGGGATTAATTTGGCTCCGGGTGAAGTGA 2678
 QY 543 TTGTTGGGATAAAGACATCGAGCTCAATCGAGGCAAGAGTAACCGAACTTTGAGTTA 602
 Db 2679 TTTTCAAAAAACGAGACATCACTTTGAATGAGGCAACCAAGCCACCACTTTAGAAAGTGC 2738
 QY 603 CTAATGAAGGGCTTAATTCCTTGCATGCTGGGTAGCCATTTCCACTTCTTTGAAGCTAACA 662
 Db 2739 ATAAACAAAGCGATCGCCCGTGAAGTGGGTCCCACTTCCCACTTCTTTGAAGTGAATA 2798
 QY 663 AGGCACATAAATTCGATCGTGAAGAGCCCTATGGCAAGCCCTAGATATTTCCCTCTGGCA 722
 Db 2799 AGCTTTTGAATTTGATCGTGAAGAGCCCTATGGCAAGCCCTAGATATTTCTCTGGAA 2858
 QY 723 ACACCTAGCAGATTGGGCGAGGACAAACCCGAAAGTGCAGTTGATTCCTCTTTGGTGGCA 782
 Db 2859 CGCTGTGCGCTTTGAACCCGCTGAGAAAAAACCCTGGAAATTTGAATTCAAATTTGGCGGTA 2918
 QY 783 GTAAAAAGTGAATGGCATGAACGGGCTTTGTGAATAACATCGCGGATGAACGCCATTAAC 842
 Db 2919 ACCAACGCATTTACGGCTTTAACTCTCTTGTGGATCGCAAGCCGATCTGATGGCAAAA 2978
 QY 843 ATAAAGCGCTTGCACAGGCGAAATCTCACGGATTTT----- 877
 Db 2979 AACTTGCTCTCAACCGCCCAAGAACATGGCTTTTGGTGTGTGAATTTGGGTGGGATGCGATA 3038
 QY 878 ----ATCAAGTAAGGAGACTCCCATG---AAAATGAAAAAACAAGAAATATGTAATACCT 930
 Db 3039 AAAAATAGGAAGGACAAATCCGATGAAGAAATCTCTCGAAAAAGAAATATGTTCTTATGT 3098
 QY 931 ACGGACCCCAACGAAGCGGATTAAGTGGCTTTAGGAGATACCGATCTTTTGGGCGAAGTAG 990

Db 3099 ATGACCCCACTACCGGCGATAAAGTGAGATTGGGCGATACCGACCTGATCTTAGAAGTCG 3158
Qy 991 AACATGACTATACCACTATCGCGAGCACTTAATTTTGGCGGGTAAACTATCCGTG 1050
Db 3159 AACATGACTGACCACTTATGCGGAAGAAATTAAGTTTGGTGGCGGTAAACCATTTCGCG 3218
Qy 1051 AGGTATGGGTGACAGCAATAGCCCTGATGAAACACCCCTAGATTAGTCTATCATCAACG 1110
Db 3219 ATGGGATGGCACAACCAACGACCCAGCAGCCAGCACTCGAICTTGTGTCTACTAAGC 3278
Qy 1111 CGATGATATGACTACACCGGGATTTACAAGCCGACATTTGGGATTAATAACCGCAAAA 1170
Db 3279 CCCTGATCGTGGATTAACACCGGCATTTATAAGCCGATATTGGCATTAATAAATGSCAAA 3338
Qy 1171 TCCATGCAATTTGGCAAGCAGGAAACAAAGCAATGCAAGATGGGTAGCCCTCATATGG 1230
Db 3339 TCCATGCAATTTGGCAAGCAGGCAATTAAGCAATGCAAGATGGGTGTCACCAATCTTT 3398
Qy 1231 TCGTGGGTGGGCACAGAGCACTAGCAGGGAGGTATGATTTATTACCGCTGGGGAA 1290
Db 3399 GCGTGGGCCCTGTACTGAGGCTTTGGCCGCTGAAGGCTGATTTACAGCTGGTGGGA 3458
Qy 1291 TCGATTCACACACCACTTCTCTCCACAACAAATTCCTACCGCTTAGCCCAATGGCG 1350
Db 3459 TTGACACCCACATCCACTTTATTTCTCCCAACAAATCCCCACAGCATTTGCCAGCGGA 3518
Qy 1351 TTACAACATGTTGGAGCGGCAAGCTCTGATGAGTGGCAGCAAGATGCTACTATCA 1410
Db 3519 TCACAACCATGATGGTGGGGAAACAGTCCAGCTGATGGGACTAAGCGGACTACCATCA 3578
Qy 1411 CTCGGGCAATGGAATTCGACCGCATGTTGCGGCGAGCAGAGAGATTTCTATGATG 1470
Db 3579 CTCGGGGCGCTGGAACCTTTAAACCAATGCTCCGTGCTCTGAAGAATATGCCATGAAT 3638
Qy 1471 TGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACTTGTAGAACAAAGTAGAG 1530
Db 3639 TGGGCTATTTGGGTAAAGGAATGTGTCTTAGAACCCCTCCCTGGTGCATCACTCGAAG 3698
Qy 1531 CGGGCGGATGTTGTTTAAATTTGATGAAGACTGGGGCAACAACCAAGTCGATCGATC 1590
Db 3699 CTGGAGCAATTTGCTTTAAATCCACGAGACTGGGTAGCACAACCTGCGCCATCTACC 3758
Qy 1591 ACTGCTTGAGCGTGGCAGATGAATACGATGCAAGTTTGTATGCCAACCGGATACAGTCA 1650
Db 3759 ATTGCTTGAATGTGGTGAACAATACGATGCAAGTGGCTATCCACACGATACCTTGA 3818
Qy 1651 ATGAGCAGGTTATGTAGATGACACCTTAATGCAATGAACGGCGCGCATCCATCCCT 1710
Db 3819 ATGAAGCGGCTGTGTGAAGACACTTTGCAAGCAATTTGCTGGGCGCACTATCCACACT 3878
Qy 1711 ACCACATTTGAGGAGGGGTGGAGGACACTCACTGATGTTATCACCATGGCAGCGGAGC 1770
Db 3879 TCCACACTGAGGTGTGTGGGGGCAACGCTCGGATGCTTAAGATGTCTGGCGAT 3938
Qy 1771 TCAATATTTACCTCTCCACACCCCACTATTCCTATACCATTAATACGGTTGCAG 1830
Db 3939 TTAACATCTCCAGCTTCTACCAACCCCACTTCTTTCCAGCTGAATACAGAGCG 3998
Qy 1831 AACACTTAGACATGCTCATGACATGCAACCACTAGCAACCGATCCGCGAGATTTAC 1890
Db 3999 AACACATGGACATGTTGATGGTGTGCCACCACTTTGGATAAAAAACATCAAAAGAGATGCC 4058
Qy 1891 AATTTTCTCAAGCCGTATCCGCGCGCTCTATCGCGGCTGAAGATGTCTCCATGATA 1950
Db 4059 AGTTTCTGATTTAGATTTCGCCCCCAACCACTTCGCCGCTGAGGACAAATCCACGATA 4118
Qy 1951 TGGGTGTGATTCGAGTGAACAAGCTCGGATTCGAAGCAATGGGCGCTGCAAGGGCAAGTGA 2010
Db 4119 TGGGATTTTCTCTATCACCAGCTCTGACTCCCAAGCGATGGGCCGTGTAGGCGAGGTCA 4178
Qy 2011 TTCCTCGAATTTGGCAGACTGCGGATGAAGATTAATAAAGAAATTTGGTAAGCTTCTGAAG 2070
Db 4179 TCACCCGCACTTTGGCAACAGCGGCAAAAAAATAAAGAAATTTGGTTCGCTTCGCTGAGG 4238

Qy 2071 ATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACC 2130
Db 4239 AAAAAAGCGGATAATGACAACTTTCCGCAATCAAGCGCTACATTTCCAAATACCACTCAACC 4298
Qy 2131 CGGCTTTGACCCACCGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATCCCG 2190
Db 4299 CGGCTATTGACACCGCATTTCTGAATATGTGGCTCTGTAGAAAGTGGGCAATTCGCG 4358
Qy 2191 ACTTGGTGGTGGAAATCTGCTTTTGGCGTAAACCCAAATCTGTATCAAAAGCG 2250
Db 4359 ATTTGGTGGTGGAGTCTGCGTTCTTTGGGCATTAACCCCAACATGATCATCAAGGCG 4418
Qy 2251 GTATGGTGGTCTTCTCTGAATGGGCGCATTTAAACCGGTCTGTGCCACTCCCAACCGG 2310
Db 4419 GATTCATCGCACTTTCTCAATGGGCGATGCCAATGCTCTATCCCACTCCCAACCGG 4478
Qy 2311 TTATTTACCGCAAAATGTTTGGGCATCACGGAAGCGGAAATTTGACACGAGATCACTT 2370
Db 4479 TGTATTTACCGCAAAATGTTTGGCCACCATGTTAAAGCCAAATTTGACCAATATCACTT 4538
Qy 2371 TTGTTTCAAAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGCCAAG 2430
Db 4539 TTGTATCCCAAGTGGCTTATGACACGCGCATTAAGAAGAGTTGGGCTTGCAAGAGTGG 4598
Qy 2431 TTCTACCGGTCAAAAACCTGCGTAACATCAACAAGAAAGACTTCAAGTTCAACGACAAA 2490
Db 4599 TTTTGGCAGTTAAAAAATCTGCCCAACATCAACAAAAAGACCTCAAAATTCACGATGTA 4658
Qy 2491 CGGCAAAATCACCGTCGATCCGAAACCTTCAGAGTCTTTGTAGATGGCAAACTCTGCA 2550
Db 4659 CGGCACATCGAAGTCAATCCTGAAACCTTAAAGTTAAAGTGGATGGCAAGAGGTTA 4718
Qy 2551 CCTTAAACCCACCTCGCAAGTGCCTTAGCCAGCGCTACACTTTCTTCTAGG 2604
Db 4719 CTTCCAAAGCAGCGGATAAATCAGCCTAGCACAACTCTACACTTGTCTTAGG 4772

RESULT 7

US-10-639-273-4
; Sequence 4, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZONII UREASE GENES AND THEIR USES
; FILE REFERENCE: IN DIAGNOSTIC AND TREATMENT METHODS
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Helicobacter bizzozeronii
US-10-639-273-4

Query Match 32.7%; Score 940.8; DB 8; Length 1710;
Best Local Similarity 72.0%; Pred. No. 3e-196;
Matches 1227; Conservative 0; Mismatches 477; Indels 0; Gaps 0;

Qy 900 AAAATGAAAACCAAGATATGTAAATACCTACGACCCCAACAAAGCGGATAAAGTCGC 959
Db 7 AAAATCTCTCGAAAAGAAATATGTTTCTATGTATGACCCACTACGGCGGATAAAGTGAGA 66
Qy 960 TTAGGATACCGATCTTTGGCGAAGTAGAATAGTACTATACCACTATCCGGAAGAA 1019
Db 67 TTGGCGGATACCGACCTGATCTTAGAGTGAACATGATGTCACCACTTATGCGGAAGAA 126

1020 CTTAAATTTGGCGGCTAAACTATCGGTGAGGATGGGTGAGCAATAGCCCTGAT 1079
127 ATTAAGTTTTGGTGGCGGCTAAACCAATTCGGGATGGGATGGCAAAACCAAGCCCAAGC 186
1080 GAAAAACACCCCTAGATTTAGTCATCACTAAACCGGATGATTAACGACTACACCGGATTTAC 1139
187 AGCCACGAACTCGATCTTGCTGCTCACTAAACCGCTGATCGGATTAACCGGATTTAT 246
1140 AAAGCCGACATTTGGATTTAAAAACGGCAAAATCATGGCATTTGGCAAGCGCAAAACAAAG 1199
247 AAAGCCGATATTGCAATTAATAATGGCAAAATCATGGCATTTGGCAAGCGCAAAATAA 306
1200 GACATGCAAGATGCGGTAAAGCCCTCATATGCTGCTGGGTGGGACAGAGCACTAGCA 1259
307 GACATGCAAGATGCGGTAAAGCCCTCATATGCTGCTGGGTGGGACAGAGCACTAGCA 366
1260 GGGGAAGGTATGATTAATACCGCTGGGGGAATGATTCACACACCACTTCTTCTCCA 1319
367 GCTGAAGGCTGATTTGACGCTGGTGGGATTTGACACCACTTCTTATTTCTCCC 426
1320 CAACAATTTCCCTACCGCTCTAGCCAAATGGCGTTTCAACCAATGTTTGGAGGCGGCACAGGT 1379
427 CAACAATTTCCCAAGCATTTGGCCAGGGGATCAACCAATGATTTGGTGGGGAAACAGGT 486
1380 CTGTGATGCGCAAGATGCGACTACTATGCTCGGGCAAAATGGAATTTGCAACGGCATG 1439
487 CCAGCTGATGGAATAAGCGGACTACCATCACTCCGGGGCGCTTGAACCTTTAAACCATG 546
1440 TTGCGCGCAGCAGAGAGTATTTCTATGAATGCTGGCTTTTGGCAAGGCAATAGCTCT 1499
547 CTGCGTCCCTCTGAAGAATATGCCATGAATGCTGGCTATTTGGTAAAGGAATGTGCT 606
1500 AGCAAAAAAACAATTTGTAGAAACAAGTAGAAGCGGCGGATTTGGTTTAAATTTGCATGAA 1559
607 TATGAACCCCTCCCTGCTGATCACTGAAAGCTGGAGCAATGGCTTTAAATCCACGAA 666
1560 GACTGGGCGCAACACCAAGTGGGATGATCACTGCTTGGAGCGGTGGGAGATGAAATACGAT 1619
667 GACTGGGCGTACACACCTCGAGCCATCTACCATGCTTGAATGTGGCTGACAAATACCAT 726
1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGGTATGTAGATGACACCCCTA 1679
727 GTGCAAGTGGCTATCCACACCGATACCTTGAATGAAGCGGCTGTGTGGAAGACACTTTG 786
1680 AATCAATGAACGCGCGCCATCCATCCCTACCACTTGAAGGAGCGGTGGAGGACAC 1739
787 CAAGCAATGCTGGGCGCACTATCCACACTTCCACTGAAGTGTGTGGGGGAC 846
1740 TCACCTGATTTATCACCATGGCAGGAGCTCAATATTTCACTCCCTCCACCAACCCCC 1799
847 GCTCCGGATGTCATTAAGATGCTGGGAAATTTACATCCTCCCAAGCTTCTACCAACCC 906
1800 ACTATTCCTATACATTAATACGGTTGCAAGACACTTAGACATGCTCATGACATGCCAC 1859
907 ACCATTCCTTTACCGTGAATACGAAGCCGCAACACATGGACATGTTGATGTGTGCCAC 966
1860 CACTTAGCAAAACGATCCGAGGATTTACAAATTTCTCAAGCGGTATCGCCCGGCG 1919
967 CACTTGGATAAAACATAAAGAGATGTCAGTTTGTGATTTAGATTTGGCCCCCA 1026
1920 TCTATCGCGGCTGAAGATGTCTCCATGATATGGGTGTGATCGGATGACAAAGCTCGGAT 1979
1027 ACCATCGCGCTGAGGACAACTCCACGATATGGGATTTTCTCTATCACCAGCTCTGAC 1086
1980 TCGCAAGCAATTTGGGCGTGGAGGAAATGATTTCTCGAACTTTGGCAGACTGGCGATTAAG 2039
1087 TCCCAAGCGATTTGGGCGGTGTGAGGAGGTCTACCCGCACTTTGGCAAAACAGCGGACAA 1146
2040 AATTAATAAGAAATTTGTAAGCTTCTCAAGATGCAAGATGCAAGATTAATTTCCGCTAT 2099
1147 AACAAAAAAGAAATTTGGTGGCTTGGTGGGAAAAAAGGCGATAATGACAACTTCCGCATC 1206
2100 AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGGGTGAGCGATAT 2159

1207 AAGCGCTACATTTCCAAATACACCATCAACCCCGCTATTGACACGGCATTTCTGAATAT 1266
2160 ATCGGCTCTGTGAGAGAGGCAAGATCGCGCACTTTGGTGGTGTGGAATCCTGCTTTTTT 2219
1267 GTCGGCTCTGTAGAGTGGCAAAATTCGCCGATTTGGTGTGTTGGAGTCTCTGCTTTCTT 1326
2220 GCGGTAAACCCCAAAATCGTGATCAAAAGCGGTATGGTGTCTTCTGAAATGGGCGAT 2279
1327 GGCATTAACCCCAACATGATCATCAAGGCGGATTCATCGCACTTTCTCAAAATGGGCGAT 1386
2280 TCTAACGCGTCTGTGCGCCACTCCCAACCGTTTATTAACCGGGAATGTTTGGGCAATCAC 2339
1387 GCCAATGCTCTATCCCCACTCCCCCAACCGGTGATTAACCGGGAATGTTTGGCCACCAT 1446
2340 GGCAGGCGAAATTTGACACCAAGCATCACTTTTGTGTTTCCAAAGTCGCTTATGAAATGGC 2399
1447 GGTAAAGCCAATTTGACACCAATATCACTTTTGTATCCCAAGTGGCTTATGACAACGGC 1506
2400 GTGAAAGAAAGCTGGGCTTAGAGCGCCAAAGTCTTACCGGTCAAAAACTGCGGTAAATC 2459
1507 ATTAAGAGAGTGGGCTTGCAAGAGTGGTGTTCGCCAGTTTAAAACTGCCCAACATC 1566
2460 ACCAAGAAAGCTTCAAGTTCACCAACGCAAAACGCAAAATCACCGTCCATCCGAAAC 2519
1567 ACCAAAAAGACCTCAAAATCAACGATGTTACCGCACATCGAAGTCAATCTCTGAAAC 1626
2520 TTGAGGCTTTGTAGATGGCAAACTCTGCACTCTTAAACCCACCTCGCAAGTGCCTCTA 2579
1627 TACAAGTTAAAGTGGATGGCAAGAGGTACTTCCAAAGCAGCGGATTAATACAGCTTA 1686
2580 GCCAGCGCTACACTTTCTCTAG 2603
1687 GCACAACTCTACAACTTGTCTAG 1710

RESULT 8
US-10-476-313-10
; Sequence 10, Application US/10476313
; Publication No. US20040241175A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, BRENT
; APPLICANT: BOYLE, JEFFREY
; APPLICANT: LEW, ANDREW
; TITLE OF INVENTION: Antigen Targeting
; FILE REFERENCE: BDWP-002
; CURRENT APPLICATION NUMBER: US/10/476,313
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PR5241
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 10
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Helicobacter felis
US-10-476-313-10

Query Match 31.3%; Score 902.4; DB 9; Length 1719;
Best Local Similarity 70.6%; Pred. No. 8.7e-188;
Matches 1203; Conservative 0; Mismatches 501; Indels 0; Gaps 0;

900 AAAATGAAAAACAAGATATGTAATACCTACGGACCCCAACCAAGCGGATAAAGTGC 959
15 AAGATTTACGAAAGATATGTTCTATGTATGTTCCCACTACCGGGATCGTGTAGA 74
960 TTAGGAGATACCGATCTTTGGGCAAGATGAACTATACCTATACCTATGCGGAAGAA 1019
75 CTCGCGCACACTGATTTGATCTTAGAAGTGGAGATGATTGCCACCACTTATGTTGGAAG 134
1020 CTTAAATTTGGCGCGGTAAACTATCCGTGAGGATGAGGTGAGCAATAGCCCTGAT 1079
135 ATCAAAATTTGGGGCGGTAAAACTATCCGTGATGGGATGAGTCAAAACCAATAGCCCTAGC 194

1080 GAAACACCCCTAGATTAGTCACTAAACCGCATGATTATCGACTACACCGGATTTAC 1139
195 TCATTATGAATTAGATTGGTGTCTCACTAAACCGCTCATTTGGAGTATACCGGATTTAC 254
1140 AAAGCCGACATTTGGATTAAAGACGGCAAAATCATGGCATTTGGCAAGGAGGAAACAAAG 1199
255 AAAGCCGACATTTGGATTAAAGACGGCAAAATTCAGGCGATTTGGCAAGGAGGAAACAAAG 314
1200 GACATGACAGATGCGGTAAGCCCTCATATGCTGCTGGGTGTTGGCAGACAGAACTAGCA 1259
315 GACATGACAGATGCGGTAAGTAATAATCTTTGGCTAGCTGCTGCTACAGAGGCTTTGGCA 374
1260 GGGAGGATGATTATACCGCTGGGGGATTCGATTCAACACCGCATTTCTTTCTCCA 1319
375 GCTGAGGCTTGTATGTAACCGCTGGTGGCATCGATACGCAATATTCATTTATCTCTCCC 434
1320 CAACAATTCCTTACCGCTCTAGCCAAATGGGTGTTAAACCATGTTTGGAGGGGACACAGGT 1379
435 CAACAATTCCTTACCGCTCTAGCCAAATGGGTGTTAAACCATGTTTGGAGGGGACACAGGA 494
1380 CTTGTAGATGGCAGATGCGATCTACTATCACTCCGGGCAAAATGGAATTTGCACCGCATG 1439
495 CTTGGATGGCAGATGCGATGCGACCATCACTCCGGGCGGCTAATCTAAAGATATG 554
1440 TTGCGGACGACAGAGATATCTAATGATGTTGGCTTTTGGGCAAGGCAATAGCTCT 1499
555 TTGCGTGCAGCGGAGAAATACGCCATGAATCTAGGCTTTTGGCTAAGGGGAAATGTGTCT 614
1500 AGCAAAAACCACTTTAGAAACAGTAAAGCGGGCGGATTTGTTTAAATTCATGAA 1559
615 TACGAACCTCTTTTACCGCATCAGATTGAAAGAGGGGCGATTTGTTTAAATCCACGAA 674
1560 GACTGGGGCAACACCAAGTGGCATCGATCACTGCTTTGAGCGTGGCAGATGAATACGAT 1619
675 GACTGGGAGACACCTGCGATTTTCAACATGCTCTCAATGTCGCGATGAATACGAT 734
1620 GTCAAGTTTGTATCCACACCGCATACAGTCAATGAGGCGGTTATGATGACACCCCTA 1679
735 GTCAAGTTGCTATCCACACCGCATACCTTTAAAGAGGGGCTGTGTAGAAGACACCCCTA 794
1680 AATGCAATGAACGGGCGGCGCATTCATGCTCAACATTTGAGGGAGGGGTGGAGGACAC 1739
795 GAGGCGATTTGCGGGGCGACCATTCATACCTTCCACATGAAAGGGGCTGGGGGTGGACAC 854
1740 TCACCTGATGTTATCCATGCGGCGAGCTCAATTTCTACCTCTCCACACACCC 1799
855 GCTCCAGATGTTATCAAAATGGCAGGGGAAATTTAAACATTTTACCCGCTCTACTAACCCG 914
1800 ACTATTCCCTATACCATTAATACGGTTGACAGAACACTTTAGACATGCTCATGACATGCCAC 1859
915 ACCATTCTTTTACCAAAAACACTGAAGCCGAGCAGCATGACATGTTAATGGTGGCCAC 974
1860 CACTAGACAAACGCAATCCGAGGATTTCAATTTCTCAAGCCGATTCGCGCCCGG 1919
975 CACTTGGATAAAAGTATCAAGGAGATGTGAGTTTGGCGATTTGCGGATTTGCGCCCAA 1034
1920 TCTATCGCGCTGAGATGCTTCCATGATATGGTGTGATCGGATGACAGCTCGGAT 1979
1035 ACTATCGCGCTGAGACCAACTTCCATGACATGGGATCTTTTCTATCAACAGCTCCGAC 1094
1980 TCGCAAGCAATGGGCGTGCAGCGGAGTATCTCGAACTTTGGCAGATTCGCGATAAG 2039
1095 TCTCAGCTATGGGACGCTGAGCGAGGTGATCACGCGACTTTGGCAGACAGCAGACAA 1154
2040 AATPAAAAAGAAATTTGGTAACTTCTGTAAGTGGCAAGATGCAACGTAATTTCCGCAAT 2099
1155 AACAAAAAGAGTTTGGGCGCTTGAAGAGAGAAAAAGCGGATTAACGAACTTTCCGCATC 1214
2100 AAGCGCTACATCTTCAAAATACATCAACCCGCTTTTGAACCCAGCGGCTGAGCGATAT 2159
1215 AAACGCTACATCTTAAATACACCATCAACCCCGCATTCGCGATGGGATTTCTGACTAT 1274

2160 ATCGCTCTGTGGAAGAGGGCAAGATCGCGACTTGTGTGTGGAACTCTGCTCTTTT 2219
1275 GTGGCTCTGTGGAAGTGGGCAAAATACGCCGACTCTGTGCTTTGGAGTCCGGCTTTCTT 1334
2220 GCGCTAAACCCCAAAATCGTGATCAAGGCGGTATGTTGGTCTTCTCTGAAATGGGCGAT 2279
1335 GCGATTAAGCCCAATATGATTAATTAAGGCGGATTTATTTGGCTCTCTCAATGGGCGAT 1394
2280 TCTAACGGTCTGTGCGCATCTCCCAACCGGTTTATTAACGCGAAATGTTTGGCATCAC 2339
1395 GCAATGGCTCTATTTCCCAACCCCTCAGCCCGTCTATTACCGTGAATGTTTGGACACAT 1454
2340 GCGAAGGCGAAATTTGACACACGATCACTTTTGTTCCTCAAGTCTGCTATGAAATGGC 2399
1455 GGGAAAAACAAATTCGACACCAATATCACTTTGCTGCCAAGCGCTTACAAGGCGAGG 1514
2400 GTGAAGAAAAAGCTCGGCTTAGAGCGCCAAAGTTCTACCGGTCAAAAACTGCCGTAAATC 2459
1515 ATCAAGAAAGACTAGGCTAGATCGCGTGTATTGCGAGTGAANAATGTCGCAATATC 1574
2460 ACCAAGAAAGACTTCAAGTTCAACGACAAACGCGCAAAATCAACCGTCAATCGAATCGAAAC 2519
1575 ACTAAAAGGACCTCAATTTCAACGATGTGACCGACATATTGATGTCAACCTGAAAC 1634
2520 TCGAGGTCTTTGTAGATGGCAACTCTGACCTCTAAACCCACCTCGAAGTGCCTCTA 2579
1635 TATAGGTGAAGTGGATGGCAAGAGGTAACTCTTAAAGCAGCAGATGAATTTGAGCCTA 1694
2580 GCCAGCGCTACACTTTCTCTAG 2603
1695 GCGCACTTATATTTGTTCTAG 1718

RESULT 9
US-10-335-977-3849
; Sequence 3849, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy B.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 3849:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1710 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...1710
SEQUENCE DESCRIPTION: SEQ ID NO: 3849:
US-10-335-977-3849

Query Match 30.8%; Score 886.4; DB 8; Length 1710;
Best Local Similarity 70.0%; Pred. No. 2.9e-184;
Matches 1193; Conservative 0; Mismatches 511; Indels 0; Gaps 0;

QY 900 A A A A T G A A A A A C A A G A A T A T G T A A A T A C C T T A C C G A C C C A A A G G C G A T A A A G T G C G 959
DB 7 A A G A T T A G C A A A A G A A T A T G T T C T A T G T A T G C C C T A C C A G G C G A T A A A G T G A G A 66

QY 960 T T A G G A G A T A C C G A T C T T T G G C A G A A G T A G A A C A T A C T A T A C C A C C T A T G G C G A A G A A 1019
DB 67 T T G G C G A T A C A G A T T G T G C T G A A G T A G A A C A T G A C T A C A C A T T T A T G G C G A G A G 126

QY 1020 C T T A A A T T T G C G C G G T A A A A C T A T C C G T A G G G T A T G G T C A G A C A A T A G C C C T G A T 1079
DB 127 C T T A A A T T C G G T G C G G T A A A A C C T A A G A A A G C A T G A G C C A T C C A A C A A C C C T A G C 186

QY 1080 G A A A C A C C T A G A T T A G T A C A T C A C T A A C C G A T A T T A T C G A C T A C A C C G G A T T T A C 1139
DB 187 A A A G A A G A A C T G G A T T G A T C A T C A T A A C G C T T T A A T C G T G G A T T A C C C G T A T T T A T 246

QY 1140 A A A G C C A C A T T G G A T T A A A A C G G A A A T C A T G G C A T T G C A A G G C A G G A A A C A A G 1139
DB 247 A A A G C G A T A T T G T A T T A A A G C G C A A A T C G T G C A T T G T A A A G G C G T A A C A A A 306

QY 1200 G A C A T G A A G A T G C G T A A G C C C T A T A T G T G T G G T G T G G C A C A A A G C A C T A G C A 1259
DB 307 G A C A T G C A A G A T G C G T T A A A A C A A T C T T A G C G T G G T C T G T A C T A G A A G C G T A G C T 366

QY 1260 G G G A A G T A T A T T A C C G T G G G G A T T C A T T C A C A C C C A C T T C T T C T C C A 1319
DB 367 G G T G A A G G C T T A A T C G T A A G C G T G T G T A T T G A C A C A C A C A T C C A C T T C A T T C A C C C 426

QY 1320 C A A C A A T T C C T A C C G T C T A G C C A A T G G C G T A C A C C A T G T T T G A G G C G G C A C A C G T 1379
DB 427 C A A C A A T C C C T A C A G C T T T G C A A G C G T A A C A C C A T G A T T G T G G C G G A A C C G G T 486

QY 1380 C C T G T A G A T G C A C A A T G C G A C T A C T A T C A C C G G C A A A T G G A A C T T G C A C C G C A T G 1439
DB 487 C C T G C T A T G C A C T A A T G C G A C T A C T A C T C C A G G T A G A A A C T T A A A T G G A T G 546

QY 1440 T T G C G C G C A G A A G A T A T T C T A T A A T G T G G C T T T T T G G G C A A A G G C A A T A G C T C T 1499
DB 547 C T C A G A G C G G C T G A A G A A T T C T A T G A A C T T A G G T T C T T A G C T A A A G G T A A C G C T T C T 606

QY 1500 A G C A A A A A C A A C T T G T A G A C A A G T A G A A G C G G C G A T T G T T T A A A T T G C A T G A A 1559
DB 607 A A C A C C G A G C T T A G C G G A T C A A A T T A A G C C G T G C G A T T G G C T T T A A A A T C C A C G A A 666

QY 1560 G A C T G G G C G C A A C A C A C C A A G T G C G A T C A T C A C T G C T T G A G C G T G C G C A G A T A A T A C C A T 1619
DB 667 G A C T G G G C A C C A C T C C T T C T G C A T C A T C A T C A T G C G T T A G A C G T T G C G A C A A T A C C A T 726

QY 1620 G T C A A G T T T G T A T C C A C C G A T A C A G T C A A T A G G C A G G T T A T G T A G A T G A C A C C C T A 1679
DB 727 G T G C A A G T G C T A T C C A C A C A G A C A C T T T A A T G A A G C C G T T G C G T G G A A G A C A C T A T G 786

QY 1680 A A T C A A T A A C G G C G G C C A T C C A T C C C T A C C A T T A G G A G C G G G T G G A G A C A C 1739
DB 787 G C A G C T A T T G C G G A C G C A C T A T G C A C A C T T T C C A C A C T A A G G C G C T G T G G C G G A C A C 846

QY 1740 T C A C C T A T G T T A T C A C C A T G C G A G C T C A A T A T T C T A C C C T C C T C C A C C A C C C C C 1799

RESULT 10

US-09-913A-251

; Sequence 251, Application US/09895913A

; Patent No. US20020160456A1

; GENERAL INFORMATION:

; APPLICANT: Kleantous, Harold

; APPLICANT: Al-Garawi, Amal

; APPLICANT: Miller, Charles

; APPLICANT: Tomb, Jean Francois

; APPLICANT: Oomen, Raymond P.

; TITLE OF INVENTION: Identification of Polynucleotides

; TITLE OF INVENTION: Encoding No. US20020160456A1 Helicobacter Polypeptides in the

; TITLE OF INVENTION: Genome

DB 847 G C T C T G A C A T T A T T A A A G T G C C G T G A A C A A C A C A T C C T A C C C G C T T C C A C T A A C C C C 906
QY 1800 A C T A T T C C C T A T A C C A T T A A T A C G G T T G C A G A A C A C T T A G A C A T G C T C A T G A C A T G C C A C 1859
DB 907 A C T A T C C C T T C C A C T G T G A T A C A G A A G C C G A A C A C A T G G A C A T G C T C A T G G T G T G C C A C 966
QY 1860 C A C T A G A C A A A C G C A T C C G G A G A T T T A C A A T T T T C A A A G C C G T A T C C G C C C C G G C 1919
DB 967 C A C T T G G A T A A A A G C A T T A A A A G A G A T G T C C A G T T C G C T G A T T C A A G G A T C C G C C C T C A A 1026
QY 1920 T C T A T C G C G C T G A G A T G T C T C A T A T A T G G G T G T G A T C C G A T G A C A G A C T C G G A T 1979
DB 1027 A C C A T T G C G C T G A A G A C A C A C T T T G C A T G A C A T G G G A T T T C T C A A T C A C C A G T T C T G A C 1086
QY 1980 T C G A A G C A A T T G G G C G T G C A G G C G A A G T A T C C T C G A A C T T G G C A G A C T G C G G A T A A G 2039
DB 1087 T C T C A A G C G A T G G G T C G T G T G G T G A G T T A C A C T A G A C T T G G C A A C A G A C A C T G A C A A 1146
QY 2040 A A T A A A A A A A T T T G T A A G C T T C C T G A A G A T G C A A A G A T A A C A G A T A A T T T C C G A T T 2099
DB 1147 A A C A A A A A A A T T T G C C G C T T G A A A G A A A A A A A G G C G A T A A C G A C A A C T T C A G A T C 1206
QY 2100 A A G C G T A C A C T C C A A A T A C A C T A T C A A C C C G C T T T G A C C C A C C G C G T G A C G G A T A T 2159
DB 1207 A A A C G C T A C T T G T C T A A A T A C A C C A T T A A C C C A G C G A T C G C T C A T G G A T T A C G G A T 1266
QY 2160 A T C G C T C T G T G A A G A G G C A A G A T C G C C G A C T T G T G T G T G G A A T C T C T G A A A T G G G C G A T 2219
DB 1267 G T A G T T C T G T A A A G T G G C A A A G T G C T G A C T T G T G T A T T G T G A G T C A G A T T C T T T 1326
QY 2220 G G C G T A A A C C C A A A T C G T A T C A A A G C G G A T G T G T G T C T T C T C T G A A A T G G G C G A T 2279
DB 1327 G G C G T A A A C C C A C A C A T A T C A T C A A A G C G G A T T C A T T G C G T T A A G C C A A A T G G G C G A T 1386
QY 2280 T C T A A C G C T G T G C C A C T C C C A A C C G G T T A T A C C G G A A T G T T G G G C A T C A C 2339
DB 1387 G C G A C G C T T C T A T C C C T A C C C C A C A C C G G T T A T T A C A G A A A T G T T C G C T C A C C A T 1446
QY 2340 G G C A A G G C G A A T T T G A C A C C A G C A T C A C T T T T G T T T C A A A G T C G C C T A T G A A A T G G C 2399
DB 1447 G G T A A G C T A A A T A C A T G A A C A C A C A C T T T T G T G T C T C A A G C G C T T A T G A A A G C 1506
QY 2400 G T G A A A A A A A G C T G G G C T T A G A G C C A A G T T C T A C C G G T C A A A A A C T G C C G T A A C A T C 2459
DB 1507 A T T A A A G A A A T T A G G C T T G A A A G A C A A G T T T G C C G T T A A A A T T G C A A A A C A T C 1566
QY 2460 A C C A A A A A G A C T T C A A G T T C A A C G A C A A A A C G G C A A A A T C A C C G T C G A T C C G A A A A C C 2519
DB 1567 A C T A A A A A A A G A C A T G C A A T T C A A T G A C A C T A C C G C T C A C A T T G A A G T C A A T C C T G A A A C T 1626
QY 2520 T T C G A G T C T T T G T A G A T G G C A A A C T C T G C A C C T C T A A A C C C A C C C T C G A A G T G C C T C T A 2579
DB 1627 T A C C A T G T T G T G T G A T G G C A A A G A A G T C A C T T C T A A A C C C A G C T A A T A A A G T G A G C T T G 1686
QY 2580 G C C C A G C G C T A C A C T T C T T C T A G 2603
DB 1687 G C T C A C T C T T T A G C A T T T C T A G 1710

FILE REFERENCE: 06132/043002
CURRENT APPLICATION NUMBER: US/09/895,913A
CURRENT FILING DATE: 2001-06-23
PRIOR APPLICATION NUMBER: US 08/881,227
PRIOR FILING DATE: 1997-06-24
NUMBER OF SEQ ID NOS: 368
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 251
LENGTH: 1815
TYPE: DNA
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: CDS
LOCATION: (51)...(1757)
US-09-895-913A-251

Query Match 30.8%; Score 885.8; DB 3; Length 1815;
Best Local Similarity 70.0%; Pred. No. 4.1e-184;
Matches 1193; Conservative 0; Mismatches 512; Indels 0; Gaps 0;

QY 900 AAATGAAAAAACAAGAAATATGTAATACCTACCGAACCCACAAAGGCGATAAAGTGGCG 959
DB 57 AAGATTAGCAAGAAAGAAATATGTTCTATGATGCGCCCTACTACAGGCGATAAAGTGA 116
QY 960 TTAGGAGATACCGATCTTTGGGCAAGAGTAGAACAATGACTATACCATATGGCGAAGAA 1019
DB 117 TTGGGCGATACAGACTTGATCGCTGAAGTAGAACAATGACTATACCATTTATGGCGAAG 176
QY 1020 CTTAAATTTGGCGGGGTAAAACTATCCGTGAGGTATGGGTGAGCAAGCAATAGCCCTGAT 1079
DB 177 CTTAAATTTGGTGGCGGTAAACCTGAGAGAGGCGATGAGCCATCCNACACCCCTAGC 236
QY 1080 GAAAAACCCCTAGATTAGTATCATCACTAACCGCGATGATTATCGACTACACCGGATTTAC 1139
DB 237 AAAGAAGAAATTTGATCTAATCATCACTAACCGCTTTAATCGTGAATTACCGGTATTAT 296
QY 1140 AAGCCGACATTTGGGATTAATAAAGCGGAAATCATGGCATTTGGCAAGGCGAGGAAACAG 1199
DB 297 AAGCGGATATTGGTATTAAGATGGCAAAATCGCTGGCATTTGTAAGGCGGTAAACAAA 356
QY 1200 GACATCAAGATCGGTAAGCCCTCATATGGTGGGTGTTGGGCAAGAGCACTAGCA 1259
DB 357 GACATCAAGATCGGTTAAACAAATCTTAGCGTAGGTCTGCTACTGAAGCCTTAGCC 416
QY 1260 GGGGAAGGTATGATTATACCGCTGGGGGAATCGATTACACACCCCACTCTCTTTCTCCA 1319
DB 417 GGTGAAGTTTGTATGTAATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 476
QY 1320 CAACAAATTCCTACCGCTTAGCCAAATGGCGTTTCAACACCATTTTGGAGCGGCGACAGGT 1379
DB 477 CAACAAATTCCTACAGCTTTTGAAGCGGTGTAACCAACCATGATTGTTGGCGGAACCTGGT 536
QY 1380 CTTGTAGTGGCAGATGCGACTACTATCACTCCGGGCAATGGAACTTGCACCGCATG 1439
DB 537 CTTGTGTATGGCACTAATGCGACTACTATCACTCCAGGCGAGAGAAATTTAAATGGATG 596
QY 1440 TTGCGCGCAGCAGAGATTTCTATGAATGTTGGCTTTTGGCAAGGCAATAGCTCT 1499
DB 597 CTCAGAGCGGCTGAAGAAATTTCTATGAATTTAGTTTCTTGGCTAAGGTTAAGCTTCT 656
QY 1500 AGCAAAAAACAATTTGAGAACAAAGTAGAGCGGCGCGATTTGGTTTTTAAATTCATGAA 1559
DB 657 AACGACGCGAGCTTAGCCGATCAATTTGAAGCTGGTGGATTTGGCTTTTAAATCCAGAA 716
QY 1560 GACTGGGCGCAACACCAAGTGGATGATCACTGTTGAGCGTGGCAGATGAATAGAT 1619
DB 717 GACTGGGCGCAACACTCTCTCTGCAATCAATCATCGTGTAGATTTGTCAGACAAATAGAT 776
QY 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGCAGGTTATGTAGATGACCCCTA 1679
DB 777 GTGCAAGTGTATCCACACAGACACTTTGAAATGAAGCCGGTTGCGTGGGAGACACTATG 836
QY 1680 AATGCAATGAACGGGCGCGCCATCCATCCCTACCACTTGAAGGAGCGGTGGAGGACAC 1739

DB 837 GCAGCTATTTCGGGAGCGCACTATGACACTTTCCACACTGAAGGTGCTGGCGGGACAC 896
QY 1740 TCACCTGATGTTATCACCATGCGAGCGCTCAATATTCTTACCTCTCCACACACCCCC 1799
DB 897 GCTCCTGATATTATTAAAGTAGCTGGTGAACAACAATTTCTCCCGCTTCCACTAACCCC 956
QY 1800 ACTATTCCCTATACCAATTAATAGGTTGAGAGCACTTAGACATGCTCATGACATGCCAC 1859
DB 957 ACTATCCCTTTTCACTGTGTAATACAGAAAGCAGAACATATGAGTATGTTGGTGTGCCAC 1016
QY 1860 CACCTAGACAAACGCGATCCGCGAGGATTTTCAAAAGCCGTATCCCGCCCGGC 1919
DB 1017 CACTTGGATATAAGCAATTAAGAAGATTTAGTTCTGCTGATTCAGAGATCCGCGCTCAA 1076
QY 1920 TCTATCGCGGCTGAAGATGCTCTCATGATATGGGTGTTGATCCGATGACAAAGCTCGGAT 1979
DB 1077 ACCATTGGCGCTGAAGACACTTTGATGATGATGAGGATTTTCTCAATCACAGCTCTGAC 1136
QY 1980 TCGAAGCAATGGGCGTGCAGGCGAAGTATCTTCGAACCTTGGGAGAGCTGGGATAAG 2039
DB 1137 TCTCAAGCTATGGGTGCTGTTGAGTATCACTAGAACTTTGGCAAAACAGCTGACAAA 1196
QY 2040 AATAAAAAGAAATTTGGTAAAGCTTCTGGAAGTGGCAAGATTAACGATATTTCCGCAAT 2099
DB 1197 AACAAAAGAAATTTGGCGCGCTTGAAGAAGAAAGGCGATTAACGAACTTCAGGATC 1256
QY 2100 AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTAT 2159
DB 1257 AAACGCTACTTCTTAATACACCATTAACCCAGCGATCGCTCATGGGATAGCGAGTAT 1316
QY 2160 ATCGCTCTGTGGAAGAGGCGAAGATCGCGACTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2219
DB 1317 GTAGTTCTGTAGAAAGTGGGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1376
QY 2220 GCGGTAAACCCCAATCGTATCAAGCGCGTATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2279
DB 1377 GCGGTAAACCCCAATCATGATCATCAAGCGCGGTTCATTGGCTTGAAGTCAAAATGGGAGT 1436
QY 2280 TCTAACGCGTCTGTGCCCACTCCCAACCGGTTCATTACCGCGAAATGTTTGGGCACTAC 2339
DB 1437 GCGAAGCTTCTTATCCCTACCCCAACACAGTTTATACAGAGAAATGTTTGGCTCATCAT 1496
QY 2340 GCGAAGCGAAATTTGACACACGACATCACTTTTGTTCCTCAAGTCCGCTATGAAAAATGGC 2399
DB 1497 GGTAAAGCCAAATACGATGCAACATCACTTTTGTCTCAAGCGCTTATGACAAAGGC 1556
QY 2400 GTGAAAGAAAAGCTGGGCTTAGAGCGCAAGTTCTTACCGGTCAAAAACCTGCGGTAAACATC 2459
DB 1557 ATTAAAGAAAGAAATAGGCGCTTGAAGACAAGTGTTCGCGGTAAAAAATTTGCAGAAACATC 1616
QY 2460 ACCAAGAAAGACTTCAAGTTCAACGCAAAAAGCGCAAAATCACCGTCCATCCGAAACCC 2519
DB 1617 ACTAAAAAGACATGCAATTAACGACACTACCGCTCAATTTGAAGTCAATCTGAAACT 1676
QY 2520 TTCGAGGTCTTTGATAGTGGCAAACTCTGCACTCTTAAACCCACCTCGCAAGTGCCTCTA 2579
DB 1677 TACCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1736
QY 2580 GCCCAGCGCTACACTTTCTCTTAGG 2604
DB 1737 GCGCAACTCTTTAGCATTTTCTTAGG 1761

RESULT 11
US-10-282-122A-22427
; Sequence 22427, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA.034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 22427
 LENGTH: 1710
 TYPE: DNA
 ORGANISM: Helicobacter pylori
 US-10-282-122A-22427

Query Match 30.7%; Score 884.8; DB 8; Length 1710;
 Best Local Similarity 70.0%; Pred. NO. 6.5e-184;
 Matches 1192; Conservative 0; Mismatches 512; Indels 0; Gaps 0;

Qy	900	AAATGAAAACAAAGATATGTAATACCTACCGAACCCACCAAGGCGATAAAGTGGC	959
Db	7	AAGATTAGCAAGAAAGATATGTTCTATGTATGGCCCTACTACAGGCGATAAAGTGGA	66
Qy	960	TTAGGAGATACCGATCTTTGGGCGAAGTAGAATGACTATACCATCTATGGCGAAGAA	1019
Db	67	TTGGCGGATACAGACTTGATCGTGAAGTAGAATGACTATACCATCTATGGCGAAGAG	126
Qy	1020	CTTAAATTTGGCGGGTAAACTATCCGTGAGGGTATGGGTGAGCAATAGCCCTGAT	1079
Db	127	CTTAAATTTGGTGGCGTTAAACCTGAGAGAGGATGAGCAATCAACACCCTAGC	186
Qy	1080	GAACAACCCCTAGATTTAGTCTATCACTAAACCGGATGATTCGACTACACCGGATTTAC	1139
Db	187	AAAGAAGAAATTTGGATCTAATCATCACTAACCGCTTTAATCGTGGATTAACCGGATTTAT	246
Qy	1140	AAACCGCATTTGGATTTAAAGAGGCGAATCCATGGCATTTGGCAAGCGAGAAACAAG	1199
Db	247	AAACCGGATTTGGATTTAAAGAGGCGAATCCATGGCATTTGGCAAGCGGTTAAAGCGGTAA	306
Qy	1200	GACATGAAGATGGCGTAAGCCCTCATATGGTCTGGTGGTGGGCAAGAGCACTAGCA	1259
Db	307	GACATGAAGATGGGTTAAAGCAATCTTAGCTAGGTCTGCTACTGAAGCCTTAGCC	366
Qy	1260	GGGAAGGTATGATTTATCCGCTGGGGGAATCGATTGACACACCACTTCCTTTCTCA	1319
Db	367	GGTGAAGGTTTGGATCGTAACCTGCTGGTGTATTTGACACACATCCACTTCATTTCA	426

Qy	1320	CAACAAATCCCTACCGCTCTAGCCAAATGGCGTTTACAAACCATGTTTGGAGCGCGCAGGT	1379
Db	427	CAACAAATCCCTACAGCTTTTTCGAAAGCGGTGTAACAAACCATGATTTGGTGGCGAAGTGG	486
Qy	1380	CCTGTAGATGGCAAGATGGCACTACTATCACTCCGGGCAAAATGGAACCTTGGACCGCATG	1439
Db	487	CCTGTAGTGGCACTAAATGGCACTACTATCACTCCAGGCGAAGAAATTTAAATGGATG	546
Qy	1440	TTGCGCGCAGCAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAGCAATAGCTCT	1499
Db	547	CTCAGAGCGGCTGAAGAAATTTCTATGAATTTAGGTTTCTTGGCTAAAGTAAGCTTCT	606
Qy	1500	AGCAAAAAACAATTTGTAGAACAGTAGAAGCGGGCGGCAATGGTTTAAATTTGCATGAA	1559
Db	607	AACGACGCGAGCTTAGCCGATCAAAATTTGAAGCTGGTGGCAATTTAAATTTCCAGAA	666
Qy	1560	GACTGGGGCACAAACCAAGTCCGATCACTGCTTGGAGCGTGGCAGATGAATACGAT	1619
Db	667	GACTGGGGCACCACTCTCTTCTGAATCAATCATGCTTAGATGTTTGCAGACAATACGAT	726
Qy	1620	GTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGGCAAGTTTATGTAGATGACACCTA	1679
Db	727	GTGCAAGTCGCTATCCACACAGACACTTTGAATGAAGCGGTTGCGTGAAGACACTATG	786
Qy	1680	AATGCAATGAAGCGGCGGCGCATCCATGCTTACCACTTGAAGGAGCGGTGGAGACAC	1739
Db	787	GCAGCTATTTGGCGGACGCACTATGCACACTTTTCCACACTGAAGGTGCTGGCGGAGCAC	846
Qy	1740	TCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTTCTACCTTCCCTCCACACCC	1799
Db	847	GCTCTGATATTTAATAAGTAGTGTGTGAACAACAATTTCTCCCGCTTCCACTAACCC	906
Qy	1800	ACTATTTCCCTATACCAATTAATACGGTTGCAGAACACTTAGACATGCTCATGACATGCC	1859
Db	907	ACTATTTCCCTTCACTGTGATACAGAACAGAACACATGACATGCTTATGGTGTGCCAC	966
Qy	1860	CACCTAGACAAACGATCGGAGGATTTACAATTTTCTCAAGCCGTATCCCGCCCGG	1919
Db	967	CACTTGGATAAAGCATTAAGAGAGATGTTTCAAGTTCGCTGATTCAGAGTCCCGCTCAA	1026
Qy	1920	TCATTCGGGCTGAAGATGCTCCATGATATGGTGTGATCCGATGACAACTCGGAT	1979
Db	1027	ACCATTTGGCGCTGAAGACACTTTTGCATGACATGGGGAATTTCTCAATCCAGCTCTGAC	1086
Qy	1980	TGCAAGCAATGGCGGCTGCGAGCGAAGTATCTCTCGAACTTTGGCAGACTCGCGGTAAG	2039
Db	1087	TCTCAAGCTATGGTGTGGTGGTGAAGTATCACTAGAACTTTGGCAACAGCTGACAAA	1146
Qy	2040	ATAAAAAAGAAATTTGGTAAGCTTCTGGAAGTGGCAAGATAACGATAATTTCCGCAT	2099
Db	1147	AAACAAAAAGAAATTTGGCGCTTGAAGAGAAAGGCGATAACGACAACTTCAGGATC	1206
Qy	2100	AAGCGTATATCTCAAAATACATATCAACCCGCTTTGACCCACCGCGTGAGCGAGTAT	2159
Db	1207	AAACGCTACTTGTCTAAATACACCATTAACCCAGCGATCGCTCATGGGATTAGCGAGTAT	1266
Qy	2160	ATCGGCTCTGCGAAGAGGCGAAGATCGCGGATTCGCGGATTTGGTGTGGGAATCTCGCTTT	2219
Db	1267	GTAGGTTCTGTAGAAGTGGGCAAGTGGCTGACTTGGTATTTGGAGTCCCGCATTTCT	1326
Qy	2220	GGCGTAAACCCCAAAATCGTGATCAAGCGGCTATGGTGGTCTTCTCTGAAATGGGCGAT	2279
Db	1327	GGCGTAAACCCCAACATGATCATCAAGCGGCTTCAATTCGCTTGAATCAATGGGTGAC	1386
Qy	2280	TCTAAGCGCTGTGGCCACTCCCAACCGGTTTATTAACCGGCAATTTGTTGGGCGATCAC	2339
Db	1387	GGCAACGCTTCTATCCCTACCCCAACCAAGTTTATTAACAGAGAAATTTTTCGCTCATCAT	1446
Qy	2340	GGCAAGGCGAATTTGACACAGCATCACTTTTGTTCCAAACTCGCTTATGAAATGGC	2399
Db	1447	GGTAAGCGCAATATCGATGCAAAACATCACTTTTGTGTCTCAAGCGGCTTATGACAAAGC	1506
Qy	2400	GTGAAAGAAAGCTGGGCTTAGAGCGCAAGTTTCTACCGGTCAAAAACTGCCGTAAATC	2459

Db	1507	ATTAAAGAAAGAAATTAGGGCTTGAAGA	CAAGTGTTCGGGTAAATAATTCGAAACATC	1566
Qy	2460	ACCAAGAAAGACTTCAAGTTCACGACAAAC	CGCAGAAATCACCCTCGATTCGAAACCC	2519
Db	1567	ACTAAAAAAGACATGCAATTCACGACATAC	CGCTCACATTGAAGTCAATCCTCGAAACT	1626
Qy	2520	TTGAGGTCCTTTGTAGATGCGAAACTCTG	CACCTCTAAACCCACCTCGCAAGTGCCTCTA	2579
Db	1627	TACCATGTGTCGTGGATGCGAAGAAAGTA	ACTTCTAAACCCAGCAATAAAGTGAGTTG	1686
Qy	2580	GCCGACGCTACACTTCTTCTTAG	2603	
Db	1687	GCACAACCTTTAGCATTTTCTAG	1710	

```

RESULT 12
; US-10-500-447A-5
; Sequence 5, Application US/10500447A
; Publication No. US20050150016A1
; GENERAL INFORMATION:
; APPLICANT: PARK, Hee-Sung
; TITLE OF INVENTION: Method for producing a recombinant protein using pollen
; FILE REFERENCE: YIOP040518US/PCT
; CURRENT APPLICATION NUMBER: US/10/500,447A
; CURRENT FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: KR 2001-71712
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; US-10-500-447A-5

```

Query Match 30.5%; Score 880; DB 10; Length 1710;
Best Local Similarity 69.8%; Pred. No. 7.5e-183;
Matches 1189; Conservative 0; Mismatches 515; Indels 0; Gaps 0;

900	QY	AAATGAAAAA	CAAGAAATATGTAATATACCTACGGA	CCCAAGCGGATAAAGTGGCG	959
7	Db	AAGATTAGCAGAA	AAAGAAATATGTTCTATGTATGCGCCCTAC	TACAGGCGATAAAGTGA	66
960	QY	TTAGGAGATACCGATCTTTGGCGAAGATGAA	CATGACTATACCACTATGGCGAAGAA	1019	
67	Db	TTGGCGATACAGACTTGATCGCTGAATGAGAA	CATGCTACCACTTTATGGCGAAG	126	
1020	QY	CTTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGT	CAGAGCAATAGCCCTGAT	1079	
127	Db	CTTAAATTTGGCGCGGTGTAAAA	CCCTAAGAGAAGCGATGAGCCCAATCTTAACAACCTTAGC	186	
1080	QY	GA AAAACACCTAGATTTAGTTCATCACCTAACCGCATGATTA	TTCGACTACACCGGGATTTAC	1139	
187	Db	AAGAAGAACTGATCTAATCATCACTAACGCTTTAATCGTGGATTA	CAACCGGTATTTAT	246	
1140	QY	AAAGCCGACATTTGGATTTAAAAACGGCAAAATCCATGGCAT	TGGCAAGGAGGAAACAAG	1199	
247	Db	AAAGCGGATATTGGTATTTAAGATGGCAAAATCGCTGGCAT	TGGTAAGGCGGTAAACAA	306	
1200	QY	GACATGCAAGATGCGGTAAAGCCCTCATATGGT	CGTGGGTGGGCAAGAGCACTAGCA	1259	
307	Db	GACATGCAAGATGCGGTAAAAACAATCTTAGCGTGGGT	CCTGCTACTGAAGCCCTTAGCC	366	
1260	QY	GGGGAAGGTATGATTTATCCCGTGGGGGAATCGATTT	CACACACCCACTTCTTTCTTCCA	1319	
367	Db	GGTGAAGGTTTGATCGTAACTGCTGGTGGTATTGA	CACACACACTTTCATCTCCCC	426	
1320	QY	CACAAATTCCTACCGCTCTAGCCATGGCGTTACACCAATG	TTTGGAGGGGCAACAGT	1379	
427	Db	CACAAATTCCTACAGCTTTTTCAGCGGTGTAA	CAACAGATGATTTGGTGGCGAACTGGC	486	
1380	QY	CTGTAGTGGCAAGATGCGACTACTATCACTCCGCGCAAA	TGGAACTTGCACCGCATG	1439	

[illegible]

Db 1567 ACTAAAAAGACATGCAATTCAACGACACTACCGCTCACATTGAAGTCAATCCTGAAACT 1626
Qy 2520 TTGAGGCTTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCCCAAGTGCCTCTA 2579
Db 1627 TACCATGTGTGCTGGATGGCAAGAGTAACTTCTAAACGAGCAATAAAGTGAGCTTG 1686
Qy 2580 GCCAGCGCTACACTTTCTTCTAG 2603
Db 1687 GCACAACCTTTTGTAGATTTTCTAG 1710

RESULT 13
US-10-476-313-11
; Sequence 11, Application US/10476313
; Publication No. US20040241175A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, BRENT
; APPLICANT: BOYLE, JEFFEREY
; APPLICANT: LEW, ANDREW
; TITLE OF INVENTION: Antigen Targeting
; FILE REFERENCE: BWP-002
; CURRENT APPLICATION NUMBER: US/10/476,313
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PRS241
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 11
; LENGTH: 1717
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-476-313-11

Query Match 30.4%; Score 876.8; DB 9; Length 1717;
Best Local Similarity 69.7%; Pred. No. 3.8e-182;
Matches 1187; Conservative 0; Mismatches 517; Indels 0; Gaps 0;

Qy 900 AAAATGAAAAAACAAGAAATATGTAATACCTACCGGACCCACCAAGGCGATAAAGTGGCG 959
Db 13 AAGATTAGCAAGAAAGAAATATGTTCTATGATGGCCCTACTACAGGTGATAAAGTGAGA 72

Qy 960 TTAGAGATACCGATCTTTGGGCGAGAGTAGAACAATGACTATACACCTATGGCGAAGAA 1019
Db 73 TTGGCGGATACAGACTTGATCGCTGAAGTAGAACAATGACTACACCATTTATGGCGAAG 132

Qy 1020 CTTAAATTTGCGCGGGTAAACTATCGTCAGGGTATGGGTGAGCAAGTACAGCAATGCCCTGAT 1079
Db 133 CTTAAATTTGCGTGGCGGTAAACCCCTAAGAGAAGCATGAGCCAATCTAAACACCTTAGC 192

Qy 1080 GAAAAACCCCTAGATTAGTCACTCACTAAACGCGATGATTATCGACTACACCGGATTTAC 1139
Db 193 AAGAAGNACTGGATCTAATCACTCACTAAACGCTTTAATCGTGATTAACCGGTATTTAT 252

Qy 1140 AAAGCGACATTTGGATTTAAAAACGGCAAAATTCATATGGCATTTGGCAAGCGAGAAACAAG 1199
Db 253 AAAGCGGATATTGTTATTAAGATGGCAAAATCGTGGCAITTTGTTAAAGCGGTAAACAAA 312

Qy 1200 GACATGAAGTGGGTAAAGCCCTCATATGTCGTGGGTGTTGGGTCACAGAGCACTAGCA 1259
Db 313 GACATGAAGATGGGTAAAGCCCTCATATGTCGTGGGTGTTGGGTCACAGAGCACTAGCC 372

Qy 1260 GGGGAAGTATGATTATTACCGCTGGGGAATCGAATTCACACACCCACTTCTCTTTCTCCA 1319
Db 373 GGTGAAGTTTGAATCGTAACTGCTGGTGGTATTGACACACATCCACTTCAATTTACCC 432

Qy 1320 CAACAATTTCCCTACCGCTCTAGCCAAATGGCGTTTACCAACATGTTTGGAGGCGGACAGGT 1379
Db 433 CAACAATTTCCCTACAGCTTTTGAAGCGGTGTAAACCAATGATTGGTGGCGGAACCTGTT 492

Qy 1380 CCTGTAGATGCCAATGCGACTACTACTACTCTCCGGGCAATGGAATTTGCAACCGCATG 1439
Db 493 CCTGTAGATGCCAATGCGACTACTACTACTCTCCAGGTAGAGAAATTTTAAATGSGATG 552

Qy 1440 TTGCGGCGAGCAAGAGTATTCTATGAATGTGGGCTTTTGGGCAAGCAATAGCTCT 1499
Db 553 CTCAGAGCGGCTGAAAGAAATATTCTATGAATTTAGGTTTCTGGCTAAAGGTAAACGTTCT 612

Qy 1500 AGCAAAAAACAACCTTTGTAGACAAGTAGAAGCGGCGCATTTGGTTTAAATTTGCATGAA 1559
Db 613 AACGATGCGAGCTTAGCCGATCAAAATTTAGGCGGTGCGATTTGGCTTTAAATTTACGAA 672

Qy 1560 GACTCGGGCACAACCAAGTGCATCATCTGCTTGAGCGTGGCAGATGAATAGCAT 1619
Db 673 GACTGGGCGACCACTCTCTTCTGCAATCAATCATGCTTAGATGTTTGGCGACAATAGCAT 732

Qy 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGTTTATGTAGATGACACCCCTA 1679
Db 733 GTGCAAGTGCCTATCCACACAGACACTTTTGAATGAAGCGGTTGTGTAGAAGACACTATG 792

Qy 1680 AATGCAATGAACGGCGGCCCATCATGCTACACCATTTGAGGAGCGGTGGAGACAC 1739
Db 793 GCTGCTATTGCTGGACGCACTATGCACACTTTCCACACTGAAGCGCTGGCGGCGGACAC 852

Qy 1740 TCACCTGATTTATCACCATGGCAGCGAGCTCAATATTCTACCCCTCTCCACACCCCCC 1799
Db 853 GCTCCTGATATTATTAAGTAGCCGCTGAACACACATTTCTCCCGCTTCCACTAACCC 912

Qy 1800 ACTATTCCTATACCATTAATACGGTTGCAGAACACTTAGACATGCTCATGACATGCCAC 1859
Db 913 ACCATCCCTTTCCACGTTGAATACAGAAAGCAGACACATGGACATGCTTATGTGTGTCAC 972

Qy 1860 CACCTAGACAAACGCGATCGCGAGGATTTACAAATTTCTCAAAAGCGGTATCCGCCCGGC 1919
Db 973 CACTTGGATAAAGCAATTAAGAAGATGTTCAAGTTTTCAGTTTCAAGGATCCGCCCTCAA 1032

Qy 1920 TCTATCGCGGCTGAAGATGCTCTCATGATATGGGTGTCATCGCGATGACAAAGCTCGCAT 1979
Db 1033 ACCATTCGCGCTGAAGACACTTTTGCATGACATGGGATTTTCTCAATCACCAGTTCTGAC 1092

Qy 1980 TCGCAAGCAATGGGCGGTGACAGCGAAGTATCTCTGAACCTTTGGCAGACTGCGGATTAAG 2039
Db 1093 TCTCAAGCGATGGGCGGTGAGGTTATCACTAGAACTTTGGCAAAACAGCTGACAAA 1152

Qy 2040 AATAAAGAAATTTGGTAAAGCTTCTGAGATGCAAGATGAACGATTAATTTCCGCAATT 2099
Db 1153 AACAAAGAAATTTGGCGGCTTGAAGAAAGAAAGGCGATAACGCAACTTCAGGATC 1212

Qy 2100 AAGCGCTACATCTCAAAATACACTATCAACCCCGCTTTCACCCACGCGGTGACGAGTAT 2159
Db 1213 AAACGCTACTGTTCTTAATACACCAATTAACCCAGCGATCGCTCATGGGACTAGGATAT 1272

Qy 2160 APTCGGCTCTGTGAAGAGGGCAAGATCGCGACTTGGTGGTGTGGAATCTTGCCCTTTTT 2219
Db 1273 GTCGGTCTGTAGAAGTGGGCAAGTAGCTGACTTGGTATTGTTGGAGTCCAGCATTTCTT 1332

Qy 2220 GCGGTAAACCCAAATCGTGATCAAGCGGATGTTGGTCTTCTCTGAAATGGGCGAT 2279
Db 1333 GCGGTGAACCTTAACATGATCATCAAGGTGGGTTCAATTCATTAAGCCCAATTTGGGCGAT 1392

Qy 2280 TCTAACGGCTCTGTCGCCACTCCCAACCGGTTTATTACGGGAAATGTTTGGGCACTAC 2339
Db 1393 GGAAGCTTTCTATCCCTACCCCTCAACCGGTTTATTACAGAGAAATGTTGCTCATCTG 1452

Qy 2340 GGCAAGGCGAAATTTGACACACGATCACTTTTGTTCAAAAGTCGCTATGAAATGGC 2399
Db 1453 GGTAAAGCTTAATACGATGCAAAACATCACTTTTGTCTCAAGCGCTTATGACAAAGGC 1512

Qy 2400 GTGAAGAAAGCTGGCTTAGCGGCTAAGTCTTACCGGTCAAAAATGCGCGCTTAACATC 2459
Db 1513 ATTAAGAAAGAAATTAGGACTTGAAGAACAAGTGTTCGCGGTAAAAAATTTGCAAAAACATC 1572

Qy 2460 ACCAAGAAAGCTTCAAGTTTCAACGCAAAAACGGCAAAATCAACCGTCGATCCGAAACC 2519
Db 1573 ACCAAAAAGACATGCAATTTCAACGACACTACCGCTCAATTTGAAGTCAATCTCTGAAACT 1632

Qy 2520 TTGAGGCTTTTGTAGATGGCAAACTCTGCACTCTCTAAACCCACCTCGCAAGTGCCTCTA 2579

Db 1633 TACCATGTTCTGCGATGCGAAGAGTAACTTCTTAACCAAGCTAATAAAGTAGCTTG 1692
Qy 2580 GCCAGCGCTACACTTCTTCTAG 2603
Db 1693 GCGCAACTCTTTAGCATTTTCTAG 1716

RESULT 14
US-09-402-100-1
; Sequence 1, Application US/09402100
; Patent No. US20010019834A1
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co, LTD
; APPLICANT: Kim, Byung-O
; APPLICANT: Shin, Sung-Seup
; APPLICANT: Yu, Young-Hyo
; APPLICANT: Park, Myung-Hwan
; APPLICANT: Choi, Deok-Joon
; APPLICANT: Jung, Hyung-Jin
; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Helicobacter
; FILE REFERENCE: 0136/0G140
; CURRENT APPLICATION NUMBER: US/09/402,100
; EARLIER FILING DATE: 1999-09-27
; EARLIER FILING DATE: KR 97-11950
; EARLIER FILING DATE: 1997-03-31
; EARLIER FILING DATE: KR 97-11951
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2385
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: ()..()
; OTHER INFORMATION: Recombinant DNA
US-09-402-100-1

Query Match 29.9%; Score 860.6; DB 3; Length 2385;
Best Local Similarity 69.8%; Pred. No. 1.7e-178;
Matches 1169; Conservative 0; Mismatches 504; Indels 1; Gaps 1;

Qy 900 AAAATGAAAAAACAAGATATGTAATACCTACGAGCCCAACCAAGGCGATAAAGTGGCG 959
Db 7 AAGATTAGCAGAAAGAAATATGCTTCTATGATGGCCCTACTACAGCGGATAAAGTAGA 66
Qy 960 TTAGGAGATACCGATCTTTGGCGAGAAAGTAGAACATGACTATACCACTATGGCGAAGAA 1019
Db 67 TTGGCGATACAGACTTGTATCGCTGAAGTAGAACATGACTATACCACTTATGTTGAAGAG 126
Qy 1020 CTTAAATTTGGCGGTTAAACTATCCGTGAGGTTATGGGTGAGCAATAGCCCTGAT 1079
Db 127 CTTAAATTTGGCGGTTAAACTATCCGTGAGGTTATGGGTGAGCAATAGCCCTGAT 186
Qy 1080 GAAACACCTAGATTAGTATCATCACTAACCGCATGATTATCGACTACCGGATTTTAC 1139
Db 187 AAAGAAGAACTGGAATTAATCATCACTAACCGCATGATTATCGGTTATACCGGTTAT 246
Qy 1140 AAAGCCGACATTTGGATTAAACCGCAAAATCCATGGCATTTGGCAAGGCGAGGAAACAA 1199
Db 247 AAAGCGATATTGTTATTAAGATGGCAAAATCGCTGGCATTTGTTAAAGGCGTTAAACAA 306
Qy 1200 GACATGCAAGATGCGTAAAGCCCTCATATAGGTGTTGGGTGGGCAAGAGCACTAGCA 1259
Db 307 GACACGCAAGATGCGTTAAACCAATCTTAGCGTGGGTCTGTACTGAAGCCCTTAGCC 366
Qy 1260 GGGGAGGTATGATTATACCGTGGGGGATCGATTACACACCCACTTCTTCTTCCA 1319
Db 367 GGTGAAGGTTTGAATGTAATCTGCTGGGTGTTATGACACACATCCACTTCTCTCCCC 426
Qy 1320 CAACAATTTCCCTACCGCTCTAGCCAAATGGCGTTTACACCATGTTTGGAGGCGGCACAGGT 1379

Db 427 CAACAAATCCCTACAGCTTTTGCAGCGGTGTAAACCAAGATTTGGTGGCGAAGTGGC 486
Qy 1380 CCGTGTAGATGGCAAGATGCGACTACTATCACTCCGGGCAAAATGGAACCTTTGCAACCGCATG 1439
Db 487 CCGTGTAGTGGCACTAAACCAACCACTATCACTCCAGGTAGAGAAATTTAAATTTCAATG 546
Qy 1440 TTGGCGGAGAGAGAGATTTCTATGAATGTTGGGCTTTTGGGCAAGCAATAGCTCT 1499
Db 547 CTCAGAGCGGCTGAAGAAATTTCTATGAATTTGGTCTTTTGGCTAAAGGTAAACGCTTCT 606
Qy 1500 AGCAAAAAACAACCTTTGTAGAACAAAGTAGAGCGGCGCATTTGGTTTAAATTTGCATGAA 1559
Db 607 AACGATGCAAGCTTAGCCGATCAAAATTTGAAGCTGGTGGCATTTGGCCCTTAAATTCACGAA 666
Qy 1560 GACTGGGCAACAAACCAAGTGCATCACTGCTTGAGCGGTGGCAGATGAATAGCAT 1619
Db 667 GACTGGGCAACCACTCTTCTGCAATCAATCATCGTGTAGATGTTGGGCAAAATAGCAT 726
Qy 1620 GTGCAAGTTTGTATCCACACCGCATACAGTCAATGAGGCAAGTTATGTAGATGACACCTTA 1679
Db 727 GTGCAAGTCTGCTATCCACACAGACACTTTGAATGAAGCGGTTGCGTGGAAAGACACTATG 786
Qy 1680 AATGCAATGAAGCGGCGCCCATCCATGCTTACCACTTGAAGCGGCTGGCGGCGGAGACAC 1739
Db 787 GACGCTATTGGCGGAGCGCACTATGACACTTACCACTGAAGCGGCTGGCGGCGGAGACAC 846
Qy 1740 TCACCTGATGTTATCCACCATGGCAGCGAGCTCAATATTTCTATACCTTCTCCACACCCCT 1799
Db 847 GCTCTGATTTATTTAAGTGGCCGCTGACACACATCTTACCCGCTTCCACTAACCCCT 906
Qy 1800 ACTATTCCTATACCAATTAATACGTTGACAGACACTTAGACATGTCTATGACATGCCAC 1859
Db 907 ACTATTCCTTTTCCCGTGAATACAGAAAGCGCAACACATGACATGCTTATGTTGTCAC 966
Qy 1860 CACTAGACAAACCGCATCCGCGAGATTTCAATTTTCTCAAGCGGTATCCGCCCGCGC 1919
Db 967 CACTTGGATAAAGCAATTAAGAGATGTCCAGTTTCGCTGATTCAAGGATTCGCCCTCAA 1026
Qy 1920 TCTATCGCGCTGAAGATGTCTCATGATATGGTGTGATCGCATGATCAAGCTCGCAT 1979
Db 1027 ACCATTGCGGCTGAAGACACATTTGATGATGATGGGATTTTCTCAATCACTAGTCTGAC 1086
Qy 1980 TCGAAGCAATGGGCGTGCAGCGAAGTGTCTCTCGAATTTGGCAGACTGCGGATAAG 2039
Db 1087 TCTCAAGCGATGGCGGTGTTGAGTTATCACTAGAACTTTGSCAAACAGCTGACAAA 1146
Qy 2040 AATAAAAGAAATTTGGTAAGCTTCTGAAAGATGGCAAGATAACGATTAATTTCCGCAAT 2099
Db 1147 AATAAAAGAAATTTGGCGCTTGAAGAAAGAAAGGCGATAACGCAACTTTTCAAGATC 1206
Qy 2100 AAGCGCTACATCTCCAAATACACTATCAACCCGCTTTTGACCCAGCGCTGAGCGAGTAT 2159
Db 1207 AAGCGCTACTTGTCTAAATACACCATTAACCCAGCGATCGCTCATGGGATTAGCGAGTAT 1266
Qy 2160 ATCGGCTCTGGAAGAGGCGAAGATCGCGACTTTGGTGTGGAATCTGCTCTTTT 2219
Db 1267 GTCGGTCTGTAGAGTGGGCAAGTGTGACTGATTTGTTATTTGGTGGTCCCGCATTTT 1326
Qy 2220 GCGGTAAACCCAAATCGTATCAAGCGCGTATGGTGGTCTTCTCTGAATTTGGCGCAT 2279
Db 1327 GGTGTGAACCCCAACATGATCATCAAGCGCGGTTTCATCGCATTTAGTCAATTTGGGTAT 1386
Qy 2280 TCTAAGCGGCTGTGCGCACTCCCAACCGGTTTATTAACCGGCAATTTGTTGGGATCAC 2339
Db 1387 GCGAAGCGCTTCTATCCCTTACCCCAACACGAGTTTATTAAGAGAAATTTGCTCTCAT 1446
Qy 2340 GCGAAGCGAAATTTGACACCGCATCACTTTTGTTCCAAAGTCGCTTATGAAATGGC 2399
Db 1447 GGTAAAGCTAATAGCATGCAAAACATCACTTTTGTCTCAAGCGGCTTATGCAAGGC 1506
Qy 2400 GTGAAGAAAGCTGGGCTTAGAGCGCAAGTTTCTACCGGTCAAAACCTCGCGTAAATC 2459

Db 563922 TATCAAGCCGATATTGGTATTCTGATGGCGGTATTGTGGGTATTGGACAGAGTAA 563863
Qy 1196 CAAGGACATGCAAGATGCGTAAGCCCTCATATAGTTCGTGGGTGGGCAAGAGCACT 1255
Db 563862 CCCTGACACCATGATGATAACGTACACCAAAATATGATTATCGGTGCAAGCAAGGATTCA 563803
Qy 1256 AGCAGGGAAGGTATGATTATACCGTGGGGGAATCGATTACACACCCACTTCTTTC 1315
Db 563802 TAACGGTGCAATTTAATTGCAACCGCTGGGTGATCGATACCCCATTCATTTATTG 563743
Qy 1316 TCCACAACAAATTCCTACCGCTTAGCCAAATGGCGTTACAAACCATGTTTGGAGGGGGCAC 1375
Db 563742 TCCACAACAGCAACAACATGCAATTTGAAGTGGCTTACACAGTTAAATGGTGGTGAAC 563683
Qy 1376 AGTCTGTAGATGGCAGCAATGGCACTATCATCTCGGGCAAAATGGAATTCACCG 1435
Db 563682 TGGCCCTGCTGATGATACACAGCAACCACTTGTACCCCTGGCGCATGATATGGAACG 563623
Qy 1436 CATGTTGGCCAGCAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAAAGCAATAG 1495
Db 563622 TATGTTTCAAGCGGCAGAAAGCCTTGGCCGGTAAACGTCGGATTTTTTGGTAAAGGCAACTG 563563
Qy 1496 CTCTAGCAAAAAACAACTTGTAGAACAAAGTAGAGCGGGCGGATTTGGTTTAAATTCGA 1555
Db 563562 TTCAACCTAGATCTCTGCGTGAACAAATTTGAAGCGGTGCAATTAGGTTTAAAAATCCA 563503
Qy 1556 TGAAGACTGGGGCAACAACCAAGTGCATGCATCACTGTTGAGCGTGGCAGATGAATA 1615
Db 563502 CGAAGACTGGGGTGCAACGCGCTGCGGTGATTGATTCGCTTTAAAGTAGCAGATGAAT 563443
Qy 1616 CGATGTGCAAGTTTGTATTCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACAC 1675
Db 563442 GGATATTTCAAGTGGCCATTTCACACAGACACGCTAAATGAAATGGGCTTTTGGAAAGACAC 563383
Qy 1676 CCTAAATGCAATGAACGGCGCGCATCCATCCCTACACATTTGAGGAGCGGGTGGAGG 1735
Db 563382 GATGAAAGCAATGATGACGAGTCAITCATCTTTCCATCGGAGGGCGAGGTGGTG 563323
Qy 1736 ACATCTCATCTGATGTTATCACCATGGCAGGCGAGCTCAATATTTACCCCTCTCCACCAC 1795
Db 563322 TCATGCACTGACATCATTTAAAGCAGCGATGATTCAACGATTATTACCTGCTTCAACCA 563263
Qy 1796 CCCCACTATTCCTATACATTAATACGGTTGCGAGAACACTTAGACATGCTCATGACATG 1855
Db 563262 CCCGACTCGTCCGTTTACCAAAAAACCACTGATGAACATTTGGATATGTTGATGGTTG 563203
Qy 1856 CCACCACTAGACAAACGATCCGCGAGATTTACAAATTTTCTCAAAGCGGTATCCGCC 1915
Db 563202 CCATCACTTAGATAAACCGGTGCGCGAAGACGTAGCTTTTGGCGATAGCCGTATCCGCC 563143
Qy 1916 CGGCTCTATCGCGCTGAAGATGTCTCCATGATATGGGTGATCGGATGACAAAGCTC 1975
Db 563142 TGAACCAATTGACGAGAGAAGATATTTTGCATGATATGGCGTCTTCTCATTTATGATTC 563083
Qy 1976 GGATTCGAAGCAATGGGGCGTGCAGGCGAAGTATTCCTCGAACTTGGCAGACTGCGGA 2035
Db 563082 AGACTCTCAAGCGATGGGACGTATTGGCGAAGTCGTTATTCTGATGCAATGGCAAACTGCAGA 563023
Qy 2036 TAAGAAATAAAAGNATTTGGTAACTCTTGAAGTGGCAAGATACGATAATTTCCG 2095
Db 563022 TAAGATGAAATGCAACGTGGTGAGCT-----AGGTAATGAAGGAAACGATAACTCCG 562969
Qy 2096 CATTAAGCGCTACATCTCCAAATACATATCAACCCCGCTTTGACCCAGCGGTGAGCGA 2155
Db 562968 TATTAAACGATATATCGGAAATACACCATCAACCCAGCAATGGACATGGTATTGGCGA 562909
Qy 2156 GTATATCGGCTCTGTGGAAGAGGGCAAGATCGCGACTTTGGTGGTGTGGAATCCTGCGCTT 2215
Db 562908 GCATATTGCTCGTTAGAAAGTGGTAAATCGCAGATATCGTGTATTGGAACCGATGTT 562849
Qy 2216 TTTTGGCGGTAAACCCCAAAATCGTGATCAAGGCGGTATGGTGGTCTTCTCTGAAATGGG 2275
Db 562848 CTTTGGCGGTAAACCTGAAGTGGTGTATTAAAAAGGCTTTATTAGCTATGCGAAAAATGGG 562789

Qy 2276 CGATTCTAACGGCTGTGTGCCCACTCCCCCAACCGGTTTATTACCGGAAATGTTTGGGCA 2335
Db 562788 CGATCCAAATGCCTCAATTTCCAACACCGCAACCTGTATTCTACCGTCCAATGTACGGTGC 562729
Qy 2336 TCACGGCAAGGCGAAATTTGACACAGCATCACTTTTGTTCCAAAGTCGCTTATGAAAA 2395
Db 562728 ACAAGGCTTAGCAACCGCAACAAACAGCAGTATTCTTTGTTTCAAGCCGCTGAAAAAGC 562669
Qy 2396 TGGCGTGAAGAAAAAGCTGGGCTTAGAGCGCCAAAGTTCTACCGGTCAAAAACTGCCGTAA 2455
Db 562668 TGATATTCTGTCGAAGTTTCGGTTTACACAAAGAACCAATTGCTGTGAAGGCTGCCGCA 562609
Qy 2456 CATCACCAAGAAAGACTTCAAGTTCAAACGACAAACGGCAAAAAATCACGTCGATCCGAA 2515
Db 562608 CGTAGGTAAAAAAGATCTGGTTTCATAATGATGTAAACACCAACATTTACTGTTGATGCTGA 562549
Qy 2516 AACCTTCGAGGCTTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGC 2575
Db 562548 ACGTTATGAAGTTTCGAGTGGACGAGAGTTAATTACCTGTGAACCAAGTGATAGCGTACC 562489
Qy 2576 TCTAGCCCGACGCTACACTTTTCTTCTA 2602
Db 562488 ATTGGGTACGCGATATTTCTTATTCTA 562462

Search completed: August 10, 2006, 03:37:28
Job time : 4674 secs

306 AAGCCATTGCTTACATTAGTG

306 AAGCCATGCTTACATTAGTGCCCATATTATGGACGAAGCGCGCCCTGGAAAAAAACCG 365

[illegible]

Db	27933	AAGTTGCAATTAGAAAGTGGTATCAACAACATATCGGTGGCGGTACAGGCGGCTCTGAGG	27999
Qy	1390	GCACGAATCGGACTACTATCACTCCGGGCAAAATGGAACTTGCACCGCATCTGTGGCGCAG	1449
Db	27993	GTACGAAGGCGCAACTGTCAACCTGGACCATTGGCACTTTACATCGTATGTGTACTCGCAG	28052
Qy	1450	CAGAAGAGTATTCTATGAATGTGGGCTTTTTGGGCAAAAGCAATAGCTCTAGCAAAAAC	1509
Db	28053	CAGAGTCATTACCGTTTGAATATTGGCTTTACTGCTAAAGGCGCAAGCTGTAAACACACAG	28112
Qy	1510	AACTTGTAGAACAAAGTAGAAGCGGGCGCATTTGGTTTTAAATTTGCATGAGACTGGGGCA	1569
Db	28113	CTTTAGTGCAAACAAATTCACGCTGGTGCAATCGGCTTAAAGGTGCATGAAAGACTGGGGTG	28172
Qy	1570	CAACACCAAGTGCAGTCGATCACTGCTCGAGCGGTGGCAGATGAATACGATGTGCAAGTTT	1629
Db	28173	CAACGCTTCAGCACTTGATCATGCTTTACAGTTCGAGATGACTATGACGTACAAATCG	28232
Qy	1630	GTATCCACACCGATCAGTCAATAGAGCGAGTTATGTAGATGACACCTTAAATGCAATGA	1689
Db	28233	CGTTGCACGCGAGATACGTTTAAATGAAGCTGGTTTTCATGGAAGAAACAAATGGCTGCAGTAA	28292
Qy	1690	ACGGGCGGCGCATCCATGCTCACCACATTTGAGGAGCGGGTGGAGACACTC3ACTCATG	1749
Db	28293	AAGATCGCGGTATGTCATATGTATCATACGGAAGGTGTGGTGGTGTGTCATGCGGCCGATT	28352
Qy	1750	TTATCACCATGCGAGCGAGCTCAATATTCTACCCCTCTCCACACCCGCCACTATTCCT	1809
Db	28353	TAATTTAAGTCTGCTGGTATGCTTAACATTTTACCTTCTTCAAACAATCCGACGTTACCTT	28412
Qy	1810	ATACCATTAAATACGGTTGCGAGAACCTTAGACATGTCTATGACATGCCACACCTAGACA	1869
Db	28413	ATACTGTAATAACAAATCGATGAACATTTAGATATGGTTATGATTAACCCATCAITTTAAATG	28472
Qy	1870	AACGCATCGCGAGGATTTACAAATTTTCTCAAAGCCGTATCGGCCCGGCTCTATCCGG	1929
Db	28473	CATCTATCCCTGAAGATATTTGGGTTTGGCGATTTCAGTATTCGTAAAGAAACCATTCGAC	28532
Qy	1930	CTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAA	1989
Db	28533	CTGAAGATGTATTTCAAGATATGGGCGTATTTAGTATGGTTAGTTCTGATTTCAACAGCA	28592
Qy	1990	TGGGCGTGCAGGCGAAGTGATTCCTCGAACTTTGGCAGACTCGGATPAAGATATAAAAG	2049
Db	28593	TGGGACGTGTCGAGAAGTAAATTACCCGTACTTTGGCAAGTTGCTCACCGTATGAAGAAC	28652
Qy	2050	AAATTTGGTAACTTCTGAAAGTGGCAAGATAACGATTAATTTCCGCATTTAAGCGCTACA	2109
Db	28653	AACGTGGATTTATAGATGGCGACAGTGAATACAAATGACAAATATCGTATTTAAGCGCTATA	28712
Qy	2110	TCCTCAAAATACACTATCAACCCCGCTTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTG	2169
Db	28713	TAGCAAAATATACAAATTAACCCAGCCATTACACATGGTATTTCTGACTATGTAGTTCAA	28772
Qy	2170	TGGAAGAGGCGAAGATCGCCGACTCGTGGTGTGGAATCTGCTCTTTTTTGGCGTAAAAAC	2229
Db	28773	TTGATGAAGGTAAATTTAGCCGACATCATTTCTTTGGGAACCAGCATTTCTTGGCGTTAAAC	28832
Qy	2230	CCAAATTCGTGATCAAAAGCGGTATGGTGTCTTCTCTGAAATGGGCGATTCTTAACCGGT	2289
Db	28833	CTGATGTCTATCGTTAAAGCGGGTTAATCAACGCTGCAATCAACGGAGATGCGAATGGCT	28892
Qy	2290	CTGTGCCCACTCCCAACCGGTTTATTTACCGCGAAATGTTTGGGCATCACGGCAAGCGCA	2349
Db	28893	CTATCCCTACTTCAGAACCTTTAAATATTCGCAAAATGTTATGTTCAATTAGTTCGAAATC	28952
Qy	2350	AAATTTGACACCGACATCACTTTTGTTTTCCAAAGTTCGCCTATGAAAATTTGGCGTGAAGAAA	2409
Db	28953	TACAAAGTACATCGATGACTTTTGTGTTTCTCAACTGCTTATGAAAACGATATTCGTAAAC	29012
Qy	2410	AGTGGGCTTAGAGCGCCGAAAGTTCTACCGGTCAAAAACCTCGCGTAAACATCCCAAGAAAG	2469
Db	29013	TTTTTAGGCTTAAACCGCAATTAAGACCTGTGCAACATATCCGTTAAATTAAGTAAAAAAG	29072


```
QY 2470 ACTTCAAGTTCAACGACAAACCGCAAAATACCGTGCATCCGAAACCTTCGAGGTCT 2529
Db 29073 ATATGAAAAACAATAATGCAACACAGATTTAGAGTTTGACCCAAACATATGAAGTTT 29132
QY 2530 TTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCCCAAGTGCTCTAGCCAGCGCT 2589
Db 29133 TTGTGTAGGAGAAAAATTAACAGCGNACCTGTCTACAGATTACCATTAACACACGCT 29192
QY 2590 ACATTTCCTCTAGG 2604
Db 29193 ATTTCTTATTCTAGG 29207

RESULT 2
US-11-348-413-11112
; Sequence 11112, Application US/11348413
; Publication No. US20060160121A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084100 (AM 101724)
; CURRENT APPLICATION NUMBER: US/11/348,413
; PRIOR FILING DATE: 2006-02-07
; PCT/US05/035471
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 11/243,445
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 60/615,573
; PRIOR FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 1276209
; SEQ ID NO 11112
; LENGTH: 1716
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1716)
; OTHER INFORMATION: WAN01UFW7; ORF: 00000000011100; Cluster contains WAN01OSH0
; OTHER INFORMATION: :SE1863:ureC:urease alpha subunit:ATCC12228:NC_004461.1
US-11-348-413-11112

Query Match 19.7%; Score 566.8; DB 9; Length 1716;
Best Local Similarity 59.4%; Pred. No. 6e-124;
Matches 1018; Conservative 10; Mismatches 672; Indels 15; Gaps 3;

QY 898 TGAATGAAAAAACAAGAAATATGTAAATACCTACGGACCCCAACCAAGCGGATGATGCTCTGA 957
Db 8 TTAATGACACAATCTCAATACACAAGTCTTTATGACCAACTGTAGGAGACTCTCTGA 67
QY 958 GCTTAGAGATACCGATCTTTTGGGAGAGATAGAACATGACTATACCACTTATGGCGAAG 1017
Db 68 GATTAGAGATACCAACTCTTTTGCAAGTTGARAAGACTATGCAAAATATGAGAGATG 127
QY 1018 AACTTAAATTTGGCGGGTAAACTATCTCGTAGGGTATGGGTGACA---GCAT--- 1070
Db 128 AGCTACTTTTCGGTGGCGGAAATCAATTCGTGATGTATGGCTCAAAATCTCTAATGTGA 187
QY 1071 --AGCCCTGATGAAAAACACCTAGATTTAGTTCATCAACGCGATCATTTATCGACTACA 1128
Db 188 CAAGATGATAAATATGATAGCCGATTTAGTTTAACTAAGCAATTAATTTATGATTATG 247
QY 1129 CCGGATTTACAAAGCCGACATTTGGGATTAATAACCGCAAAATCCATGGCATTGGCAAG 1188
Db 248 ACAAGATTGTTAAAGCAGAVATCGGAATTAATAATGGTTATATTTTAAAGATCGTAAAG 307
QY 1189 CAGAAACCAAGGACATGCAAGATGGCGTAAAGCCCTCATATGGTGTGGGTGGGACAG 1248
Db 308 CTGGAACCCAGATATAATATGGATAAGCT-----TGACATCATCATTTGGTGCAACACTG 361
```

```
QY 1249 AAGCACTAGCAGGGGAAGGTATGATTTATTACCGTGGGGGAATCGATTACACACCCACT 1308
Db 362 ATATTATTGCTGCTGAGGTAAAAATTTGTTACTGCGCGGTATCGATACACACGTGCACT 421
QY 1309 TCTCTTTCCCAACAATTTCCCTACCGCTCTAGCCAAATGGCGTTTACAACCATGTTTGGAG 1368
Db 422 TVATCAATCTCTGAAACAAGCTGAAGTTGCACTTGAGAGTGGTATTACAACGCATATCGGTG 481
QY 1369 GCGGCACAGGTCTCTGTAGATGGCAAGATGCGACTATCATCTCCGGGCAAAATGGAACCT 1428
Db 482 GAGGAACTGGTCTTCTGAAGGTGCTTAAAGCGACTACTGTAAACACCGAGCACTTGGCATA 541
QY 1429 TGCACCGCATGTTGCGCGCAGCAGAGAGTATTCTATGATGATGGGCTTTTGGGCAAG 1488
Db 542 TTCATCGCATGTTAGAACGAGCAGAGAGATGCTTATTAATGTAGGATTTACTGGTAAAG 601
QY 1489 GCAATAGCTCTAGCAAAAAACAATTTGTAGAAACAAGTAGAAGCGGCGCATTTGGTTTTA 1548
Db 602 GKCAAGCTGTCAATCATACTGCACTTTTGAACAAATTCATGCGAGGCTATAGTCTTAA 661
QY 1549 AATTGCATGAAGACTGGGGCAACAACAAGTGGCATCGATCACTGCTTGAGCGTGGCAG 1608
Db 662 AAGTACATGAAGATTTGGGAGCTACACCTTCAGCATTAAGTCAATGCAATTAGAGCTTGCAG 721
QY 1609 ATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCAAGTTATGTAG 1668
Db 722 ATGAGTTTGTATTTCAAGTCGCTTTACATGCGAGACATTAATAAGAGCTGGGATTTATGG 781
QY 1669 ATGACACCTTAAATGCAATGAACGGCGCGCCATCCATCGCTACCACTTACCATTTGAGGAGCGG 1728
Db 782 AAGTACAAATGGTCTGCTGTAAGATCGTGTATTGCAATGATGATCATCTAAGAGGAGCTG 841
QY 1729 GTGGAGGACACTCACCTGATGTTATCACCATGGCGGCGAGCTCAATATTTCTACCTCTCT 1788
Db 842 GTGGTGGTCACTGCACCTGACTTAATCAAAATCAGTGCATATTCAAACATCTTACTTCTT 901
QY 1789 CCACCAACCCCACTATTTCCCTATACCATTAATACGGTTGCGAGAACTATTAGACATGCTCA 1848
Db 902 CTACAAACCCCACTATTTACTTTACACKCAACAACACTGTAGTAGAACATTTAGACATGTTA 961
QY 1849 TGACATGCCACCACTAGACAAACCGCATCCGCGAGGATTTACAATTTTCTCAAGCGCTA 1908
Db 962 TGATVACTCACCATCTTAAATGCTTCAATACCAAGAGACATTCGATTTGAGATTTCTCGTA 1021
QY 1909 TCCGCCCCGGCTCTATCGCGCTGAAGATGTGCTCCATCATATGGGTGTGATCGCGATGA 1968
Db 1022 TAGTGAAGAACTATAGCAGCAGAGAGATTAATACAGATATGGCGTATTAGTATGG 1081
QY 1969 CAAGCTCGGATTCGCAAGCAATGGGCGCTGCAAGCGAAGTGAATTCCTCGAACTTTGCAGA 2028
Db 1082 TAAGTTCAAGATTCAAGCAATGGGACGTGTCGGTGAAGTTGTAAACACGCTACTTGGCAAG 1141
QY 2029 CTGCGGATGAANTAAAGAAATTTGGTAAGTTCCTGAAGATGCGCAAGTGAAGATAAGATA 2088
Db 1142 TTGCACACCGTATGAAGAAACAACCGCGACCAATAGATGGTGATTTGAATATCACGATA 1201
QY 2089 ATTTCCGCTTAAGCGCTACATCTCCAATACACTATCAACCCCGCTTTTGACCCCGCG 2148
Db 1202 ATATCGTATTAAAGTTCAATTTGCAAAATATACAATCAATCTGCGCATATACATGGTA 1261
QY 2149 TGAGCGATATATCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTTGGGTGTGGAATC 2208
Db 1262 TTTCTGACTATGTTGGATCTGTGAAGCGGTAAACTTTGCCGATTTAGTAAATGTGGGAAC 1321
QY 2209 CTGCTTTTGGCGTAAACCCAAATCGTGATCAAAAGCGGTATGGGTGCTTCTCTG 2268
Db 1322 CAGAATTTCTCGGTGCCAAACCCGATCTTGTGTTTAAAGTGGCATGATTAATCTWGCAG 1381
QY 2269 AAATGGCGATTTCTAAACCGGTCTGTGCCCACTCCCCAACCGGTTTATTAATCCGCGAATGT 2328
Db 1382 TAAATGGTGTGCTAAATGGCTCCATACCAATCAGAGCTTTTGAATATCGCAAAATGT 1441
QY 2329 TTGGGCATCACGGCAAGGCGAAATTTTGACACGAGCATCACTTTTGTGTTTCCAAAGTCGCT 2388
```

Db 1442 ATGGTCAATTTGGTGTAAACATTTACACATACCTGCTATGACTTTTGTCTTAAACACTGCAT 1501
Qy 2389 ATGAAATATGCGGTGAAGAAAGCTGGGCTTAGAGCGCCAAAGTTCTACCGGTCAAAAAC 2448
Db 1502 ATGAAATATGCGGTGAAGAAAGCTGGGCTTAGAGCGCCAAAGTTCTACCGGTCAAAAAC 1561
Qy 2449 GCGGTAAACATCACCAAGAAAGCTTCAAGTTCAACGCAAAACGGCAAAAATCACCGTGC 2508
Db 1562 TTAGAAATTTAACTAAGGCAGATATGAAAAATAAATATGCTACACCTTAAATATAGATGAG 1621
Qy 2509 ATCCGAAACCTTCGAGGTCTTTGTAGATGCGAACTCTGCACCTCTTAAACCCACCTGCG 2568
Db 1622 ATCCGAAACATATGAGGTATTCGTTGATGGTATTAATAATCAAGTGAAGCAGACACAG 1681
Qy 2569 AAGTGCTCTAGCCAGCGCTACACTTTCTTCTAG 2603
Db 1682 AATTACCWTTAACACAAAGATACTTCTTATCTAG 1716

RESULT 3
US-10-471-571A-3299
; Sequence 3299, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471, 571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3299
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-3299

Query Match 18.7%; Score 539.6; DB 6; Length 1713;
Best Local Similarity 58.9%; Pred. No. 1.6e-117;
Matches 1010; Conservative 0; Mismatches 689; Indels 15; Gaps 4;

Qy 896 CATGAAATCAAAAACAAGATATGTAATACCTACGCGACCAACCAAGCGGATAAAGT 955
Db 6 CTTTAAATGACGCAAAATCAATATACGCTTATACGGTCCAACTGTTGGAGATTCAT 65
Qy 956 GCGCTTAGGAGATACCGATCTTTGGGAGAGAGTAAATACATGATATACCACTATGGCGA 1015
Db 66 TCGTTTAGGTGATACGATCTATTTGCTCAATAGAAAAGACTATGCGGTTTATGGTGA 125
Qy 1016 AGAAGCTTAAATTTGGCGGGTAAATATCCGTTAGGGTATGGTCA-----GAGC 1067
Db 126 AGAAGCTTAAATTTGGTGGTAAATCTATAGAGACGGTATGGCGCAAAATCCCTCGTGT 185
Qy 1068 AATA-GCCCTGATGAACACCCCTAGATTTAGTTCATCACTAAACCGGATGATTCGACTA 1126
Db 186 AACACGTGATGACGTGAACCGTTGACAGACCTTGTCATTTCTAAATGCGGTTATTCGATTA 245
Qy 1127 CACCGGATTTACAAAGCGACATTTGGATTTAAACCGCAAAATCCATGCGATTGGCAA 1186
Db 246 CGATAAGTGGTTAAAGCTGATATAGGCAATTAATAATGGTTATATTTTCGCCATAGGTAA 305
Qy 1187 GGCAGGAAACAAGGACATGCAAGATGCGGTAAAGCCCTCATATGGTTCGTGGGTGGGCAAC 1246
Db 306 TGCCGGCAACCCGATAT-----AATGATATGTC--GACATTAATATAGGTTCAACAAC 359
Qy 1247 AGAAGCACTAGCAGGGAAGGTATGATTTATACCGCTGGGGGAATCGATTACACACCCA 1306
Db 360 AGATATCAITGGCCCTGAAGGTAAATTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
Qy 1307 CTTCCCTTTCCCAACAATTTCCCTCTACCGCTCTAGCCCAATGGCGCTTACAACCATGTTGG 1366

Db 420 TTTTATTAAATCTGAAACAAGCAGAGGTGCGCATTAGAAAGTGGTATTACGACTCATATTGG 479
Qy 1367 AGCGGCGACAGGTCTCTGTAGATGGCAGAAATGCGACTACTACTCATCTCCGGGCAAAATGGAA 1426
Db 480 TGGTGGTACTGGTCTTCAGAAAGTTCTTAAGCAACAACACTGTAACTCCAGGTCCATGGCA 539
Qy 1427 CTTGCAACCGCATGTCGCGCAGCAAGAGAGTATTCTATGAATGGGGCTTTTGGGCAA 1486
Db 540 TATTATAGAAATGTTAGAAAGCTGCCGAAGTTTACCGATTAATGTCGGTTCGTACAGGTAA 599
Qy 1487 AGGCAATAGCTCTAGCAAAAACAACTTTGTAGAACAAAGTAGAGCGGGCGCATTTGGTTT 1546
Db 600 AGGCAACAAACAATCCAACTGCACTCATTTGAACAAAATCAATGCGGAGCAATTGGATT 659
Qy 1547 TAAATTCATGAAGACTGGGCGCAACAACCAAGTGGCATCGCATCTGCTGTCAGCGTGGC 1606
Db 660 AAAAGTATAGAGACTGGGGTGCACACCATCTGCTTTGAGTCTGATCATGATAGATTGTC 719
Qy 1607 AGATGAATACGATGTGCAAGTTTGTATCCACAACCGATACAGTCAATAGAGCGAGTTATGT 1666
Db 720 TGATGAATTTGATGTTCAAAATTTGCATTTACATGACAGATACTTTTAAATGAAGCAGGATTTAT 779
Qy 1667 AGATGACACCCCTAAATGCAATGAACGGCGCGCCATCCATGCTTACCACATTTGAGGAGC 1726
Db 780 GGAAGACACAATGGCTGCTGTTAAAGACCGGTGACTTCTATATGTACCATACTGAAGGTGC 839
Qy 1727 GGGTGGAGGACACTCACCTGATGTTTATCACCATGGCAGCGAGCTCAATATTTCTACCCCTC 1786
Db 840 TGGTGGCGGTCATGGGCTGATTTAAATCCGCTGCAATTTTCAATATTTTACCTTC 899
Qy 1787 CTCACACACCCCACTATTTCCCTATPACCATTAATACGGTTGAGAACACTTACAGATGCT 1846
Db 900 ATCTAGAAATCCAACTTTTGGCTTATACATAATACTGTAGATGAACATTTAGATATGTT 959
Qy 1847 CATGACATGCCACACCTAGACAAACGGATTCGCGAGGATTTACAAATTTTCTCAAGCCG 1906
Db 960 AATGATTAATCACTCACTTTTAAATGCGGCTATTCTGAAGATATCGCATTCGCAATTCACG 1019
Qy 1907 TATCGCCCGGCTCTATCGCGGCTGAAGATGCTGCTCCATGATATGGGTGATCGCAT 1966
Db 1020 TATTCGTAAGAAACGATTTGACGAGAGAGATGTTCTGCAAGATATGGGTGATTCAGTAT 1079
Qy 1967 GACAAGCTCGGATTCGCAAGCAATGGGCGTGCAGCGCAAGTGAATTCCTCGAACTTGGCA 2026
Db 1080 GATTAGTTCCGATTCACAAGCAATGGGCGTGTAGTGAAGTAAATTACACGACATGGCA 1139
Qy 2027 GACTCGGATGAAGATAAATAAAGATTTGGTAAGCTTCTGAAAGATGGCAAGATAAACGA 2086
Db 1140 AGTAGCACATCGCATGAAAGAAACAACGTTGCTTTAGATGGTGAATTTTGAACATATGA 1199
Qy 2087 TAAATTCGCGATTAAGCGCTACATCTCCAAATACACTATCAACCCGCTTTGACCCACCG 2146
Db 1200 TAAATTCGCGATCAACCGTTATATCGTAAATATACAAATTAACCCGCAATTAACATGG 1259
Qy 2147 CGTAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTGGTGGGA 2206
Db 1260 TATTTCTGATATGTAGGATCTATCGAGCGGGCAAACTAGCTGACATTTGCTTATGGGA 1319
Qy 2207 TCCTGSCCTTTTGGCGTAAACACCAAAATCGTGAATCAAAAGCGGTATGGTGGTCTTCTC 2266
Db 1320 CCCAATTTCTTTGGGGTTAAACCTGAATTAGTTGTAAAGGGCGGATTAATTAACCTCTGC 1379
Qy 2267 TGAATGGCGGATTTCAACGGCTCTGTGCCACTCCCAACCGGTTTATTTACCGCGAAT 2326
Db 1380 CGTAAATGCGGATGCAAAATGGTTCTTATACCTACATCTGAACCGATGAAGTACCGTAAAT 1439
Qy 2327 GTTTGGGATCAACGGCAAGGGGAAATTTGACACACGACATCACTTTTGTTCCAAAGTCG 2386
Db 1440 GTATGTCATATACGGCGGAAACCTTACAAGTAGGTCAATGACATTCGTGTCTTAAACTGC 1499
Qy 2387 CTATGAAATGCGGTGAAAGAAAAGCTGGGCTTTAGAGCGCCCAAGTTCTACCGGTCAAAA 2446

Db	1500	TTATGAAAAATGGTATCAATCGTGCATTAAATTTTAAAAACGATGGTGGTCCAGTTAAAAA	1559
Qy	2447	CTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCACACGACGACAAAACGCAAAATTCACCGT	2506
Db	1560	TATTAGACAATTATCTTAAGCAGNATGAAAAATAACAGTGCACACTAAATTAGAGCT	1619
Qy	2507	CGATCCGAAAAACCTTCGAGGTCCTTTGTAGATGGCAAACTCTGCACCTCTTAAACCCACCTC	2566
Db	1620	TGATCCCAACAACATATCAAGTATATGTAGATGGAGAAAAAATTACAAGTAAATGCAGCAAC	1679
Qy	2567	GCAAGTCGCTCTAGCCGAGCGCTACACTTTTCCTC	2600
Db	1680	TGAGTTACCAATTAACTCAAGAGATCTCTTTATTTC	1713

RESULT 4

US-11-216-545-652
; Sequence 652, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: Mc Laird, Paul L
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Soybeans
; FILE REFERENCE: 38-21 (53659)B
; CURRENT APPLICATION NUMBER: US/11/216,545
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 652
; LENGTH: 2673
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (588)..(588)
; OTHER INFORMATION: n is a, c, g, or t
US-11-216-545-652

Query Match	18.4%;	Score 528.6;	DB 8;	Length 2673;
Best Local Similarity	57.2%;	Pred. No. 7.7e-115;		
Matches 978;	Conservative 0;	Mismatches 729;	Indels 3;	Gaps 1;
QY	911	ACAAGAAATGTAATACCTACCGACCCCAAGGCGATAAAGTCGGCTTAGAGATAC	970	
Db	822	AGAGGAATATGCTCAACAAGTATGCCCTCAACTGGTGACAAAATCCGCTTCGGTGATAC	881	
QY	971	CGATCTTTGGGCGAAGTAGAAATGACATATACCACTATGCGCGAAGAACTTAAATTTGG	1030	
Db	882	TGACTTTGTTGCTGAATTTGAAAAAGATTTTGCTCTCTATGGTGATGATGCTGTTTTGG	941	
QY	1031	CGCGGTAAACATATCCGTGAGGTTATGGTCAAGACAAATAG---CCCTGATGAAACAC	1087	
Db	942	AGGTGGGAAAGTTTAAAGAGATGGAATGGGTCAATGATGGGGATCCACCTCSAATCTC	1001	
QY	1088	CCTAGATTTTAGTCATCACTAAACCGGATGATATTCGACTACACCGGGATTTACAAAGCCGA	1147	
Db	1002	CTTTGGATCTGTTATTAACAAATGCTCTGATTTTCGGCTCGTGGTGGGAATCATCAAGCGGA	1061	
QY	1148	CATTGGGATTAACAAACGGCAAAATCCATGGCATTTGCCAGGACGAGAAACAAGACATGCA	1207	
Db	1062	TATAGGCATTTAAAGATGGTCTCATTTGTTTCAATTTGGAAAAGCAGGAAATCCAGACATCAT	1121	
QY	1208	AGATGGCGTAAGCCCTCATATGTTCTGTGGTGTGGGCACAGAAGCACTAGCAGGGGAAGG	1267	
Db	1122	GGATGATGATATTTTTTAATATGATCAATGCGGCTAATATCTGAAGTTATTGCTGGAGAGG	1181	
QY	1268	TATGATTATTACCGCTGGGGGAATCGATTTCACACACCCACTTCTCTTCTCCACAAATTT	1327	

Db	1182	GTTGATTGTAAACAGCAGGGGCTATAGATTTGTCACGTGCATTATATATATGCCCCTCAATTAGT	1214
Qy	1328	CCCTACCGCTCTAGCCCAATGGCGTTACAACCATGTTTGGAGCGGCACAGGTCCTGTAGA	1387
Db	1242	AGATGAGGCCATATCAAGCGGCATCAACAATTAGTGGAGGTGGAAACCGGACCGACTGC	1301
Qy	1388	TGGCAGCAATGCGACTACTATCATCTCGGGGCAAAATGGAACTTGCACCGCATGTTTGGCGGC	1447
Db	1302	TGGAAACAGTGGCCACAACCTTGTAACACGACCACTCTCAGATGAAACTTAATGCTGCAATC	1361
Qy	1448	AGCAGAGAGTATTCTATGATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTAGCCAAAA	1507
Db	1362	AACTGATGACCTGCCTCTAAACTTTGGTTTCACTGGCAAAAGGAAGTAGTTCCAAGCCCTGA	1421
Qy	1508	ACAACTTGTAGAACAAAGTAGAAGCGCGCGATTTGTTTAAATTGCATGAAGACTGGGG	1567
Db	1422	TGAGTCCCTCGAGCCGATTCGGCTCGAGCAATGGGACTGAAGCTGCATGAGGACTGGG	1481
Qy	1568	CACAACCAAGTGGCATCGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGT	1627
Db	1482	AAGTACTCCCGCTGCAATAGACAGTGTCTTGACTGTTGCTGATCAATACGATATCCAGAT	1541
Qy	1628	TTGTTCCACACCGATACAGTCAATGAGGCAGGTATGTAGATGCACCCCTTAATATGCAT	1687
Db	1542	TAACATATACACCAACCCCTAAATGAAGCTGGAAATTTGTGCAACATAGCATTTGCAGCAT	1601
Qy	1688	GAAACGGCGGCCCATCCATGCTTACCACATTTAGGGAGCGGTGGAGGACACTCACCTGA	1747
Db	1602	TAAAGGAAGAACTATTCTATCTTACCAAGTGAAGGTGCAGGTGGTGGTCACTGCTCCAGA	1661
Qy	1748	TGTTTATCACCATGGCAGCGAGCTCAATATTCTACCTCTCCACCAACCCCACTATTTC	1807
Db	1662	TATCATCAAGATATGTGTTATGAAGATGTTTGCCTCATCAACAACCCCAACACGCC	1721
Qy	1808	CTATACCAATTAAATAGGTTGCAGAACTCTTAGACATGCTCATGACATGCCACCACTAGA	1867
Db	1722	TTTAACTCTCAATACTATAGATGAGCATCTTGACATGTTGATGGTCTGCCATCATCTGAA	1781
Qy	1868	CAAAACGATCCGGAGGATTTACAAATTTTCTCAAAGCGTATCCGCCCGCTCTATTCG	1927
Db	1782	TAGGGAAATTCAGAAAGCTTAGCTTTTGATGTTCAAGATTAAGAAAGAACGATTC	1841
Qy	1928	GGCTGAAGATGTGCTCCATGATATGGGTGTGATCGGATGCAAGCTCGGATTCGCAAGC	1987
Db	1842	TGCCGAAGATATTTTGATCATATATGGGCAATTAGCATCATATCTTCTGATTCCTCAGC	1901
Qy	1988	AATGGGCGGTGCAGGCGAAGTGAATCTCGAACTTTGGCAGACTGCGGATTAAGAAATAAAA	2047
Db	1902	TATGGGTGCAAGTTGGAGAGGTGATAGACAACTTTGGCAAACTTGCCAAATGAAGTAGAGT	1961
Qy	2048	AGAAATTTGGTAAGCTTCTCGAAGATGCAAGATAAGATAATTTCCGCAATTAAAGCGCTA	2107
Db	1962	ACAACGAGAACCATCCAGCCTGGTGAATCCGAAATGACAACTTCGCGATCAAAAGATA	2021
Qy	2108	CATCTCCAAATACACTATCAACCCCGCTTTGACCCAGCGCGTGAGCGAGTATATCGGCTC	2167
Db	2022	CATTGCAAAATACATATAATCCAGCTATAGCGAATGGTTTTTTCAAATATGTTGGTTC	2081
Qy	2168	TGTGGAAGAGGGCAAGATCCCGCACTTGTGGTGTGGAATCTCGCTCTTTTGGCGTAAA	2227
Db	2082	GGTGGAGGTGGGAAAAGTTAGTCTGATCTTGATGATGGAAGCCATCTTTTTTTGGGGCAAA	2141
Qy	2228	ACCCAAATCGTATCAAGGCGGTATGGGTGCTTCTCTGAAATGGCCATTTCTAACGC	2287
Db	2142	ACCAGAAATGGTATCAAGGTGGGTGGTTGCATGGGCTGATATGGGTGACCCAAATGC	2201
Qy	2288	GTCTGTGCCCACTCCCCAACCCGGTTTTATTACC CGGAAATGTTTTGGGCATCACGCGAAGGC	2347
Db	2202	AAGCATCCCCACTCTCTGAACCGGTGAAGTAGGCGCTATGTTTGGAACTACTAGGCAAGC	2261
Qy	2348	GAAATTGACACAGCATCATCTTTGTTTCCAAAGTGCCTCTATGAAATGGCGTGAAGA	2407
Db	2262	TGTTGGTGTCTTTATTCATAGCTTTTGTGACGAGGAGGTGTAGACACAGAGAGTACATGC	2321

```
Qy 2408 AAAGCTGGGCTTAGAGCGCCCAAGTTCTACCGGTCAAAAACCTGCCGTAAACATCAACCAAGAA 2467
Db 2322 TCTATACGGACTGAACAAGAGGGTGGGAAGCAGTAGGCAATGTGAGGAAGCTCACTAAACT 2381
Qy 2468 AGACTTCAAGTTCAACACAAAACGGCAAAATCACCGTGCATCGGAACCTTCGAGGT 2527
Db 2382 AGACATGAAACTTAATGACTCTCTTCCACAATCACTGTGCGACCCAGATAACTACCTGT 2441
Qy 2528 CTTTGTAGATGGCAAACTCTGCACTCTTAACACCCACCTCGCAAGTGCCTCTAGCCAGCG 2587
Db 2442 TACAGCAGATGGCAGGTTCTCACAGTTTGGCAACCACTTTGTTCCCTTTCTCGAA 2501
Qy 2588 CTACACTTCTCTTAGGCACAAATGCCCT 2617
Db 2502 TTACTTCTCTTTAAGTACTAGTCCCAT 2531

RESULT 5
US-10-449-902-18794
; Sequence 18794, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18794
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK069176
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-18794

Query Match 17.7%; Score 509.6; DB 6; Length 2709;
Best Local Similarity 58.7%; Pred. No. 2.4e-110;
Matches 1007; Conservative 0; Mismatches 684; Indels 24; Gaps 5;

Qy 916 AATATGTAAATACCTACGGACCCCAAGGGGATAAAGTGGCTTAGAGATACCGATC 975
Db 849 AATATGAAGCATGTATGGGGCTCAACTGTGTGCAAGATTAGACTTGGTGACACTGACC 908
Qy 976 TTTGGGCAGAGTAGAATCACTATACCTATCGGCAAGAACTTAAATTTGGCGCG 1035
Db 909 TTTTGGCGAGATTGAAGAAGCACTATGCCATCTATGTTGATGATGTCATATTTGGAGCG 968
Qy 1036 GTAAAACTATCCGTGAGGTTATGGTTCAG--AGCA-ATAGCCCTGATGAAACACCCCTAG 1092
Db 969 GAAAGGTTCTACGTGATGGAATGGGACAGTCCGAGGGGTATCCAGCCTCAGATTGCCCTG 1028
Qy 1093 ATTTAGTCATCACTAAACCGATGATTATCGACTACACCGGATTTACAAAGCCGACATTG 1152
Db 1029 ACACAGTTGTAAACAATGCTCTTGTGATTGACTATACAGGAATATACAAAGCTGACATTG 1088
Qy 1153 GGATTAATAAAGCGCAAAATCCATGTCATTGGCAAGCGAGGAAAC---AAGGACATGCA-- 1207
Db 1089 GTATAAATGGTGGACTTATTCGTTGCTATTTGGAGGCTTGGAAACCCCTGATGTCATGGACA 1148
Qy 1208 -AGATGGCGTAAAGCCCTCATATGTTGCGTGGTGGGACAGAAAGCACTAGCAGGGGAAG 1266
Db 1149 TGGATGTTGTTAATGAAGAGATGATTGTGGGGTTTAACTGAACTTATTCGCTGAAG 1208
```

```
Qy 1267 GTATGATTATTACCGCTGGGGAATCGAATTCACACACCACCTTCCTTTCTTCCACAACAAT 1326
Db 1209 GCATGATTGTTTACTGCTGCTGGGAATTGATTGCGCATTTCACTTCATATGCGCTCAGTTGG 1268
Qy 1327 TCCCTACCGCTCTAGCCTAATGGCGTTTACAAACATGTTTGGAGCGCGCACAGGTCCTGTAG 1386
Db 1269 CAGAAGAGGCAATTTGCAAGTGGCATCAACAACATTTGGTGGGAGGTGGAACCTGGACCGAC 1328
Qy 1387 ATGGCACAAGATGCGACTACTATCACTCCGGGCAAAATGGAACTTGCACCGCATGTTGCGG 1446
Db 1329 ATGGAACCTTGGCACAACCTTGCACCCCTTACCATCTCACAATGAAACTAATCTACAGT 1388
Qy 1447 CAGCAGAAGAGTATTCTATGAATTTGGGCTTTTGGGCAAAAGGCAATAGCTCTTAGCAAAA 1506
Db 1389 CCACCTGATGAATTACCAATCAATATGGGATTCACAGGCAAGGGAATACTACAAAACCTG 1448
Qy 1507 AACAACTTTGTAGAACAAAGTAGAGCGGGCGCATTTGGTGTAAAATTTGCATGAAGACTGG 1566
Db 1449 ATGGATTGGCTGAGATCATTTAAGGCGAGGCAATGGGCTTGAAGTTTGCATGAAGATTGG 1508
Qy 1567 GCACAACACCAAGTCCGATCGATCACTGCTTCAGCGTGCAGATGAATACGATGTGCAAG 1626
Db 1509 GAAGCACCCTAGCTCGGATAGATAACTGTTATCTGTCGAGAGCTTTTGTGATATCAGG 1568
Qy 1627 TTTGTATCCACACCGATACAGTCAATAGGCGAGGTTTATGTAGATGACACCTTAAATGCAA 1686
Db 1569 TAAATATCCACACAGACACCTTAAATGATGAGTCAGGCTGTGTGGAGCATACAAATGCGCAT 1628
Qy 1687 TGAACGGGCGCGCATCCATGCTTACCACATTTGAGGAGCGGGTGGAGGACACTACCTG 1746
Db 1629 TTAAGATAGGACAAATACATATCATATAGTAGTGAAGGAGCGGTGGTGTCTAGTCCAG 1688
Qy 1747 ATGTTATCACCATGCGAGCGAGCTCAATATTCTACCTCTCCACACCCCACTATTC 1806
Db 1689 ACATTATCAAAGTATGTGGAGTGAAAAATGTTACCTCTCTTCAAACAATCCAACCTCGGC 1748
Qy 1807 CCTATACCAATTAATACGGTTGCAGAACACTTGAACATGCTCATGACATGCCACCACTAG 1866
Db 1749 CATTTACTTTGNAACACTGTAGATGAGCACCTTGTGATGCTGATGGTCTGCCATCACCCTG 1808
Qy 1867 ACAACGCAATCCGAGGATTTACAAATTTTCTCAAAGCGTATCCGCGCGGCTCTATCG 1926
Db 1809 ATAGAAATATCCAGAAGATGTAGCATTTGCCGAGTCTAGAAATTCGAGCTGAAACAATAG 1868
Qy 1927 CGCTGGAAGATGTCTCCATGATATGGGTGTGATCCGATGACAAAGCTCGGATTCGCAAG 1986
Db 1869 CTGCTGAGGATATCTTGCATGACATGGGAGCAATCAGTATCATATCATCCGATTCGCGAG 1928
Qy 1987 CAATGGGCGGTGCGAGCGAAGTGATTCTCGAACTTTGGCAGACTGCGGATAAGAAATAAAA 2046
Db 1929 CTATGGGCGCATTTGAGAGGTGATTCTCGACATGGCAAACTGCAATTAAGATGAGA 1988
Qy 2047 AAGAAATTTGGTAAGCTTCTCT-----GAAGATGGCAAGATAACGATAATT 2091
Db 1989 GACAGAGGTTAGATTACCTATATCCAGTTTCCCGCATGCTGCAGAGGCAATGACAATT 2048
Qy 2092 TCCGCAATTAAGCGCTACATCTCCAATACACTATCAACCCCGCTTTGACCCACCGGCTGA 2151
Db 2049 TCCGATTAAGAGATACATAGCTAAATACCAATAAATCTTCTATAGTGAATGGCTTTT 2108
Qy 2152 GCGAGTATATCGGCTCTGTGGGAAGGCGCAAGATCCCGCACTTTGGTGTGTGGAATCCCTG 2211
Db 2109 CTGATTTTGTGGCTCTGTGGAGTGGGAAATTAGCTGACCTTGTATTATTTGGAACCTT 2168
Qy 2212 CTTTTTTGGCGTAAACCCCAAAATCGTGATCAAAAGCGGTATGGTGGTCTTCTCTGAAA 2271
Db 2169 CTTTCTTTGGACAAAACCCAGAAATGGTTATAAAGGTGTGTCTATTGTCATGCGCTAATA 2228
Qy 2272 TGGGCGATTTCAACGGCTGTGCGCACTCCCAACCGGTTTATTACCGGGAATGTTG 2331
Db 2229 TGGGTGATCCAAATGTAGCATTTCCAACACCTGAAACCGGTTATGATGCGCACCTATGTTG 2288
```

Qy 2332 GGCATCAGCGCAAGCGAAATTTTGACACGAGCATCATTCTTTTCCAAAGTCGCTATG 2391
Db 2289 GTGCATTTGGAGGGCTGGGAAGTGCCTATCTATAGCAATTTGTGAGCAAGGCTGCTAAG 2348
Qy 2392 AAAATGGCGTGAAGAAAGCTGGGCTTTAGAGCGCAAGTTTCTACCGGTCAAAACTGCC 2451
Db 2349 AAGCTGGGTTGCAGTGCAGTACAAAGCTGGGAAAGAGGTTGGAAGCTGTAGTTCGGTTTC 2408
Qy 2452 GTAACATCACCAAGAAAGACTTCAAGTTCAACGCAAAACGGCAAAAATCACCGTCGATC 2511
Db 2409 GTGGTTTGACAAAGCTGAATATGAAACTTAAACGATGCACCTTCCGAAAAATTTGATGTCGATC 2468
Qy 2512 CGAAACCTTCGAGGCTCTTCTAGATGCAAACTCTGACCTCTTAAACCCACCTCGCAAG 2571
Db 2469 CTGAACCTTACACGGTTACTGCTGATGAGAGGTTTTCAGATGTCAACCAACCCACAG 2528
Qy 2572 TGCCTCTAGCCAGCGGTACACCTTTCTTCTTAGGCA 2606
Db 2529 TACCACATCTCGGAATTAATCTTCTCTTTAGACA 2563

RESULT 6

US-11-218-305-14860
; Sequence 14860: Application US/11218305
; Publication No. US20060141495A1

GENERAL INFORMATION:

; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: Tao, Nengbing
; APPLICANT: McLaird, Paul L.
; APPLICANT: Wu, Kunsheng

TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping

; TITLE OF INVENTION: Corn.

; FILE REFERENCE: 38-21 (53660)B

; CURRENT APPLICATION NUMBER: US/11/218,305

; PRIOR FILING DATE: 2005-09-01

; PRIOR APPLICATION NUMBER: US 60/606,880

; NUMBER OF SEQ ID NOS: 25043

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 14860

; LENGTH: 3120

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (16)..(16)

; OTHER INFORMATION: n is a, c, g, or t

US-11-218-305-14860

Query Match 17.2%; Score 495.4; DB 9; Length 3120;
Best Local Similarity 57.6%; Pred. No. 5.8e-107;
Matches 982; Conservative 0; Mismatches 706; Indels 18; Gaps 3;

Qy 916 AATATGTAATACCTACGACCCCAAGCGGCAATAAGTCGCTTAGAGATACCGATC 975
Db 1083 AGTATGCGAGCAATTATGAGCACTACCAACCGGTGATAAGATTAGGCTTGGTGATACCAATC 1142
Qy 976 TTTGGCAGAGTAGACATGACTATACCACTATGCGGAGAACTTAAATTTGGCGGG 1035
Db 1143 TTTATGCTGAGATTGAAGAGGACTTCGCCCTTTTATGTTGATGAGTGCAATATTTGGCGGTG 1202
Qy 1036 GTAAACATATCCGTGAGGGTATGGGTC-AGAGCAATAG--CCCTGATGAAACACCCCTAG 1092
Db 1203 GAAAAGTTCTCGGTGATGGCATGGGACAAGCTTCAGGGTACCCAGATCTTCTGCGCTAG 1262
Qy 1093 ATTATGTCATCACTAACCGGATGATTATCGACTACACCGGGATTTACAAAGCCGACATTG 1152
Db 1263 ATACAGTTATAACAAATGCTGTTGTGATTGATTATACCGGAATATACAAGGCTGATATCG 1322
Qy 1153 GGAATTAACCGCAAAATCCATGGCATTTGCGAAGGAGCAAAACAGGACATGCAAGATG 1212
Db 1323 GTATAAAGGTGAGCTTATAGTTGCTATTGGAAGGCTGGAAACCCCTGATGTCATGGATG 1382

Qy 1213 GCGTAAGCCCTCATATGGTCTGTGGGTGGGCACAGACACTAGCAGGGGAGGTATGA 1272
Db 1383 GTGTCCATAAACCACTGATTTGTGGGGTCAATACTGAAGTTATTGCAATGGAAGCATGA 1442
Qy 1273 TTATTACCCCTCGGGGAATCGAATTCACACCCCACTTCTCTTCTCCACAACAATTCCTTA 1332
Db 1443 TTGTCACTGCTGGTGGCAGTAGATTGCCATGTCCACTTTATATGCTCTCAGTTGGCAGAAG 1502
Qy 1333 CGGCTCTAGCCAAATGGCGTTTAAACCATGTTTGGAGCGGCGACAGGTCCTGTAGATGGCA 1392
Db 1503 AGGCAATTGCAAGCGGCATCAACGCAATTTGGTGGTGGAAACCGGACCAAGCATGGAA 1562
Qy 1393 CGAATGCGGACTACTATCACTCCGGGCAATGGAACCTTGCACCCCATGTTTGGGCGCAGAG 1452
Db 1563 CTTGTGCCCAAACTTGCACCTCCAGCACCAATCTCAATTGAAATTAATTTGTCAGCTCCAATG 1622
Qy 1453 AAGAGTTATTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCTTAGCAAAAAACAAC 1512
Db 1623 ATCAATTGCCAATTAACTGGGANTTCAAGGCAAGGGAATTAATTTCAAAACCTGAAGGAT 1682
Qy 1513 TTGTAGAACAGTAGAAGCGGCGGATTTGTTTAAATTTGCATGAAGACTCGGGGCAAA 1572
Db 1683 TGGCTGAATCATTAAGAGCTGGAGCAATGGGTTTGAAGCTGCATGAAGATTGGGGAACATA 1742
Qy 1573 CACCAAGTGGGATCGATCACTGCTTGGAGCGTGGCAGATGAATAGATGTGCAAGTTTGTGA 1632
Db 1743 CCCCATCCGCGATAGATAATTTGTTATCTGTGTGCAAGAGATTTTGTATATTTCAGGTCAAATA 1802
Qy 1633 TCCACACCGATACAGTCAATGAGGAGGATTATGTAGATGACACCCCTAAATGCAATGAACG 1692
Db 1803 TACACACAGATACCTTAAATGAATCAGGCTGCGTGGAGCATACATCGCAGCTTTTAAAG 1862
Qy 1693 GCGCGCCATCCATGCTACCACTTTGAGGAGCGGCTGGAGGACACTCACTCCCTGATGTTTA 1752
Db 1863 GTAGAGCCATACATACATATCACAGTGAAGTGCTGTGGCGGTGATGCTCCAGACATCA 1922
Qy 1753 TCACCATGGGAGCGAGCTCAATATTTCTACCTCTCTCCACACCCCACTATTTCCCTATA 1812
Db 1923 TCAGAGTTTGTGGGGTAAATAATGTTGCTCTTTCACAAATAATCTACCCGCCATTTTA 1982
Qy 1813 CCATTAAATACGGTTTCAGAACACTTTAGACATGCTCATGACATGCCACCACTAGACAAAC 1872
Db 1983 CTTGGAACACTGTAGACGAGCACCTTGATATGCTGATGGTTTGTACCACTTTGATAAAA 2042
Qy 1873 GCATCCGCGAGGATTACAAATTTTCTCAAAGCGGTATCGCCCGGCTCTATCGCGCTG 1932
Db 2043 ACATTCAGAGATGTAGCAATTTGCGGATCTAGAAATTCAGGTGAAACAAATTTGCTGCTG 2102
Qy 1933 AAGATGTCTCCATGATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGG 1992
Db 2103 AGGACATATTGCATGACATGGAGCCATAAGTATTATATCATCTGATTCACAGGCCATGG 2162
Qy 1993 GCGGTGAGGCGAAGTGTCTCTCGAACTTGGCAGACTCGGGATGAGATTAATAAAGAAAT 2052
Db 2163 GCGCGTTTGGAGAGGTGTAAACGAACTGGCAAACTGCAAAACAGATGAAGGTACAA 2222
Qy 2053 TTGGTAAAGCTTCTCGAAGATGCAAGATAA-----CGATAATTTCCGCA 2097
Db 2223 GAGGTGTTTACCTGGATCTGTGACGCTAATCGCCCGCCGACAGTGAACAACCTCCGTA 2282
Qy 2098 TTAAGCGCTACATCTCCAAATACATATCAACCCCGCTTTTGACCCCAAGCGGTGAGCGAGT 2157
Db 2283 TAAGAGATACATAGCAAAATACAAATATCCAGCAATAGTGAACGGGTTTTCAGACT 2342
Qy 2158 ATATCGGCTCTGTGGAAGAGGCAAGATCGCGCACTTGGTGTGTGGAAATCTCTGCCCTTTT 2217
Db 2343 TTGTTGGTTCTGTTGAGGTGGGAAATTAGCTGACCTTTGTTCTTTGGAAACCAATCTTTCT 2402
Qy 2218 TTGGGTAAACCCCAAAATCGTGAATGCAAGGGGTATGGTGTCTCTCTGAAATGGGG 2277
Db 2403 TCGGCGCTTAAACCCAGAACTGGTTGTAAGGGAGGTTGCAATTCATGCGGTGTAACATGGGTG 2462
Qy 2278 ATTCTAACGCGTCTGTGCCCACTCTCCCAACCGGTTTATTATCCGCGAAATGTTTGGGCATC 2337

```
Db 2463 ATCCCAATGCTAGCATTTCCAAACCTGAACCTGTTGTGATCGACCTATGTTTGGTGCAT 2522
Qy 2338 ACGCAAGGCGAAATTTGACACCGACATCACTTTTGTTCAAAGTCCGCTATGAAATG 2397
Db 2523 TTGGAAGGCTGGAAGTTCCAAATTCATTTGATTTGTGAGCAAGGCTGCTAAGAGCTG 2582
Qy 2398 GCGTGAAGAAAGCTGGGCTTAGAGCGCCCAAGTTCTACCGGTCAAAACCTGCCGTAAACA 2457
Db 2583 GTGTGCAACCGGATACAGACTAGAAAAGAGGGTGAAGCTGTAGGGCGTGTTCGAGGCC 2642
Qy 2458 TCACCAAGAAAGACTTCAAGTTCAACGACAAAACGCGCAAAATCACCGTCGATCCGAAA 2517
Db 2643 TGACAAAGCTTGACATGAAGCTCAACGACGCCCTCCCGAAAATAGAGGTCCACCGGAGA 2702
Qy 2518 CTTTCGAGGCTTTTGTAGATGGCAAACTCTGCACCTCTAAACCCCACTTCGCAAGTGCCTC 2577
Db 2703 CCTACACAGTGACCGCTGATGAGAGGTTTGTGACATGCCAACCCACGACCTGCAC 2762
Qy 2578 TAGCCACGCGCTACACTTTCTTCTAG 2603
Db 2763 TGTCTCGGAATTACTTCTCTTCTAG 2788
```

RESULT 7

```
US-10-953-349-33824
; Sequence 33824, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33824
; LENGTH: 2912
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33824
```

```
Query Match 17.1%; Score 493.8; DB 6; Length 2912;
Best Local Similarity 57.5%; Pred. No. 1.4e-106;
Matches 981; Conservative 0; Mismatches 707; Indels 18; Gaps 3;

Qy 916 AATATGTAATACCTACGACCCACCAAGCGGATAAAGTGGCTTAGGAGATACCGATC 975
Db 1067 AGTATGCGAGCATTTATGACCTACACCGGTGTAAGATTAGGCTTTGGTGATACCAATC 1126
Qy 976 TTTGGCGAGAGTAGAATCATCATATACCACTATGGCGGAACCTTAAATTTGGCGGG 1035
Db 1127 TTTATGCTGAGATTGAAAAGGACCTTCGCTTTTATGGTGATGAGTGCAATTTGGCGGTG 1186
Qy 1036 GTAAAACTATCCGTGAGGATGTGGTC-AGAGCAATAG--CCCTCATGAAAACACCCCTAG 1092
Db 1187 GAAAAGTTCTCGGTGATGCGATGGCAAGGCTTCAGGGTACCCGAAATCTTCTGCTTAG 1246
Qy 1093 ATTTAGTCACTAATACGCGATGATATTCGACTACACCGGGATTTACAAAGCCGACATG 1152
Db 1247 ATACAGTTATAACAAATGCTGTGTGATTGATTATACCGGAATATACAAAGGCTGATATCG 1306
Qy 1153 GGATTAAAAACCGCAAAATCCATGCTATGGCATTTGGCAAGGAGAAACAGACATGCAAGATG 1212
Db 1307 GTATAAAAGGTGGACTTTATAGTTGCTATTTGAAAAGGCTGGAAACCCCTGATGTCATGGATG 1366
Qy 1213 GCGTAAGCCCTCATATGCTGCTGGTGTGGGACACAGAACACTAGCAGGGGAAGGTATGA 1272
Db 1367 GTGTCCATAACAAATGATTTGGGGTCAATACACTGAAAGTTATTGCACTGAAAGCATGA 1426
Qy 1273 TTATTACCGCTGGGGGAATCGATTACACACCACTTCCTTTCTCCACAAACAATTCCTTA 1332
```

```
Db 1427 TTGTCACTGCTGGTGGCATAGATTGCCATGTCCACTTTATATGTCTCAGTTGGCAGAG 1486
Qy 1333 CGCTCTAGCCCAATGCGCTTACAAACATGTTTGGAGGCGGCACAGGTCTCTGTAGATGGCA 1392
Db 1487 AGGCAATTTGCAAGCGGCATCACGACATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1546
Qy 1393 CGAATGCGACTACTATCACTCGGCGCAATGGAATTGTCACCGCATGTTGGCGGCGAGCAG 1452
Db 1547 CTTGTGCCACAACCTTGCACCTCCAGCACCACTCTCAATTTGAAATTTAATGTTGCACTG 1606
Qy 1453 AAGAGTATTTCTATGAATGTGGGCTTTTGGGCAAGAGCAATAGCTCTAGCAAAAAACAAC 1512
Db 1607 ATCAATTTGCCAATTAACATGGGATTCACAGGCAAGGAAATACTTCAAAACCTGAGGAT 1566
Qy 1513 TTGTAGAAACAAGTAGAAGCGGCGCATTTGGTTTAAATTTGCATGAAGACTGGGGCAAA 1572
Db 1667 TGGCTGAAATCAATTAAGCTGGAGCAATGGGTTTGAAGCTGCATGAAGATTGGGGAAC 1726
Qy 1573 CACCAAGTGGCATCGATCACTCTTGGGCTGGCAGATGAATACGATGTGCAAGTTTGTGA 1632
Db 1727 CCCCATCCGCGATAGATAATTTGTTATCTGTGTCAGAAAGATTTTGATATTTCAAGTCA 1786
Qy 1633 TCCACACCGATACAGCTCAATGAGGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAAC 1692
Db 1787 TACACACAGATACCTTAATGAATCAGGCTGCTGGAGCATACATCGCAGCTTTTAAAG 1846
Qy 1693 GCGCGCCATCCATGCTTACCAACATTTGAGGAGCGGCTGGAGGACATCACTCACTGATGTTA 1752
Db 1847 GTAGAGCCATACATACATATCACAGTGAAGGTGCTGGTGGCGTCACTGCCAGACATCA 1906
Qy 1753 TCACCATGGAGGCGAGCTCAATATTTACCTCTCTCCACCCACCTATTTCCCTATA 1812
Db 1907 TCAAAAGTTTGTGGGGTAAAAAATGTGTGCCCTCTTTCACAAATCTTACCGGCCATTTA 1966
Qy 1813 CCATTTAATACGCTTCGAGAACACTTAGACATGCTCATGATGACATGCCACCACTAGACAAAC 1872
Db 1967 CTTGGAACACTGTAGACGAGCACCTTGATATCTGATGTTGTGACCACTCTGATATAA 2026
Qy 1873 GCATCCGCGAGGATTTACAAATTTTCTCAAAAGCGGTATCGCCCCGGCTCTATCGCGCTG 1932
Db 2027 ACATTTCCAGAGATGTAGCAATTTGCGAGTCTAGAAATTCGAGCTGAAACAAATTTGCTG 2086
Qy 1933 AAGATGTGCTCATGATATGGGTGTGATCGGATGACAGCTCGGATTTGCAAGCAATGG 1992
Db 2087 AGGCATATTTGATGACATGGGAGCCATAAGTATTTATATCATCTGATTTCAAGGCCATGG 2146
Qy 1993 GCGCTGCGAGGCAAGTATTCCTCGAATTTGCGAGACTGCGGATAAGATAAAGAAAT 2052
Db 2147 GCGCGTTTGGAGAGGTGATAACACGAAACATGGCAAACTGCAAAACAGATGAAGGTACAA 2206
Qy 2053 TTGGTAAAGCTTCTCGAAGATGGCAAGATAA-----CGATAAATTTCCGCA 2097
Db 2207 GAGGTAGTTTACTGGAATCTGGTACGCTAATGCGCCCGGACAGTGACAACTCGTA 2266
Qy 2098 TTAAGCGCTTACATCTCCAATAACATATCAACCCCGCTTTGACCCACGGGCTGAGCGAGT 2157
Db 2267 TAAAGAAGATACATAGCAAAATACAAATAAATCCAGCAATAGTGAATGGGTTTTCAGACT 2326
Qy 2158 ATATCGGCTCTGTGNAAGGCGCAAGATCGCGACTTGGTGGTGGTGGTGGTGGTGGTGGT 2217
Db 2327 TTTGTGGTTCTGTGAGGTGGGAAAATTTAGCTGACCTTGTCTTTGGAAACCATCTTTCT 2386
Qy 2218 TTGGGCTTAAACCCAAAATCGTGTATCAAAAGCGGTATGGTGGTCTTCTCTGAAATGGGCG 2277
Db 2387 TCGGGCTTAACACAGNACTGGTTGTAAAGGAGGTGCAATTCATGGGCTTAAATGGGTG 2446
Qy 2278 ATTTCAACGCGCTGTGCGCACTCCCAACCCGGTTTATTACCGCGAAATTTGGGGCATC 2337
Db 2447 ATCCCAATGCTAGCATTCCAACACCTGAACTGTTGTGATGCGACCTATGTTGGTGCAT 2506
Qy 2338 ACGGCAAGCGGAAATTTGACACCGCATCACTTTTGTGTTTCCAAAGTCGCTATGAAATG 2397
Db 2507 TTTGGAAGGCTGGAAGTTCCAAATTCAAATTTGTCATTTGTGAGCAAGGCTGTAAAGAGCTG 2566
```

Qy	2398	CGCTGAAGAAAAGCTGGGCTTAGAGCGCCAGTTCACCGTCAAAAACTGCGGTAAACA	2457
Db	2567	GTGTGCGCAACGGAGTACAGACTAGAAAAGAGGGTGGAAAGCTGTAGGCGGTGTTTCGAGGCC	2626
Qy	2458	TCACCAGAAAGACTTCAGTTCAACACAAAACGGCAAAAATCACCGTCGATCCGAAAA	2517
Db	2627	TGACAAAGCTTGACATGAAGCTCAACACGCGCTCCCGAAAAATAGAGGTGCAACCCGAGA	2686
Qy	2518	CGTTCGAGGTCCTTTGTATAGTGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCCTC	2577
Db	2687	CCTACACAGTGCAGCGCTGATGGAGAGGTTTTGACATGCCAACCAAGCACCACACTGCCAC	2746
Qy	2578	TAGCCAGCGCTACACTTTCTTCTAG	2603
Db	2747	TGCTCGGAATTACTTCCTGTTCTAG	2772

RESULT 8

```

US-11-292-431-36
; Sequence 36, Application US/11292431
; Publication No. US20060121061A1
GENERAL INFORMATION:
; APPLICANT: COLE, Garry T.
; APPLICANT: CHEN, Xia
; APPLICANT: SESHAN, Kalpathi R.
; APPLICANT: HUNG, Chiung-Yu
; APPLICANT: XUE, Jianmin
; APPLICANT: YU, Jieh-Juen
; TITLE OF INVENTION: Attenuated Vaccine Useful for
; FILE REFERENCE: Immunizations Against Coccidiosis spp. Infections
; FILE REFERENCE: 529522000600
; CURRENT APPLICATION NUMBER: US/11/292,431
; CURRENT FILING DATE: 2005-12-02
; PRIOR APPLICATION NUMBER: US 60/633,399
; PRIOR FILING DATE: 2004-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36.
; LENGTH: 5588
; TYPE: DNA
; ORGANISM: Coccidiosis posadasii
US-11-292-431-36

```

Query Match	16.8%;	Score 483.4;	DB 8;	Length 5588;
Best Local Similarity	60.3%;	Pred. No. 5.1e-104;		
Matches 832;	Conservative 0;	Mismatches 541;	Indels 6;	Gaps 2;
QY	902	AATGAAAAACAAGATATGTAAATATCTACGACCCCAAGAGCGCATAAAGTGCCTTT	961	
DB	2636	AATGACGCGGAAGCATATGCTCGAATGTTGGTCTCTCCACTGGAGATGTAGTCAAGCT	2695	
QY	962	AGGAGATACCGATCTTTGGCAGAGTAGAACAATGACTATACCACTATGTCGCGAAGACT	1021	
DB	2696	AGGAACACAGAGATTTCTGGATTAAAGTCGAAAAGGACCTGACCTACTATGGTGACGAATG	2755	
QY	1022	TAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGTCA---	1078	
DB	2756	TTCAATTCGGTGGCAAGACCATAGAGACGCGGATGGGGCAGCTACAGGAAGGCATTC	2815	
QY	1079	TGAAACACCCTAGATTTAGTTCATCATCACTACCGCATGATTATCGACTACACCGGATTTA	1138	
DB	2816	CGTGGATGTCCTGGATACAGTCTCGTGTGAACCGCTAAATTTGTCGATTTGACCGGTATTTA	2875	
QY	1139	CRAAGCCGACATTTGGGATTAAAAAACGGCAAAATCCATGGCATTTGCCAAGCGAGGAACAA	1198	
DB	2876	CRAAGCTGATATTGGACTATAAAGATGGATTGATCTGCGGAATCGGCAAAAGCTGGAAACCC	2935	
QY	1199	GGACATGCAAGATGGCGTAAAGCCCTCATATGGTCTGTGGGTGTGGGCACAGAAGCACTAGC	1258	
DB	2936	AGCATGATGGATGGTGTACCCCCCAACATGATAGTTGGTCTCTTCGACAGATGTTATGCG	2995	
QY	1259	AGGGGAAGGTATGATTATTACCGCTGGGGGAATCGATTACACACCCACTTCTCTTCTCGC	1318	

RESULT 9
US-11-348-413-11113/c
; Sequence 11113, Application US/11348413
; Publication No. US20060160121A1
; GENERAL INFORMATION:


```
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084100 (AM 101724)
; CURRENT APPLICATION NUMBER: US/11/348,413
; CURRENT FILING DATE: 2006-02-07
; PRIOR APPLICATION NUMBER: PCT/US05/035471
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 11/243,445
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 60/615,573
; PRIOR FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 1276209
; SEQ ID NO 11113
; LENGTH: 315
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(315)
; OTHER INFORMATION: WANO1UFWB; ORF: 0000000011100; Cluster contains WANO1UOT1
; OTHER INFORMATION: :egad_360118_2391::EGAD orf prediction:ATCC12228.NC_004461
; OTHER INFORMATION: .1
US-11-348-413-11113

Query Match      3.3%; Score 94.4; DB 9; Length 315;
Best Local Similarity 58.9%; Pred. No. 1.3e-12;
Matches 188; Conservative 2; Mismatches 123; Indels 6; Gaps 1;

Qy 1114 TGATTATCGACTACACCGGGATTTACAAAGCCGACATTGGGATTTAAACCGCAAAATCC 1173
Db 314 TAATTATTGATTATGACAAGATTCTTAAAGCAGAVATCGGAATTTAAATAATGTTATATTT 255

Qy 1174 ATGCCATTGGCAAGCGGAGAAACAGGACATGCAAGATGGCGTAAGCCCTCATATGTCG 1233
Db 254 TTAAGATCGGTAAAGCTTGGAAACCCAGATATAATATGGATAACGT-----TGACATCATCA 201

Qy 1234 TGGGTGTGGGCACAGAAGCACTAGCAGGGGAAGGTATGATTATTACGCTCGGGGAATCG 1293
Db 200 TTGGTGCACACACTGATATTAATCTGCTGAAGGTAAATTTGTTACTGCCGGCGGTATCG 141

Qy 1294 ATTCACACACCCACTTCCTTTCTCCACAACTTCCCTACCGCTCTAGCCAAATGGCGTTA 1353
Db 140 ATACACAGTGCACCTTATCAATCTGAACAAGCTGAAGTTGCACTTGAGAGTGTATTA 81

Qy 1354 CAACCATGTTTGGAGCGGCACAGGTCTCTAGATGGCAGAAATGGCACTACTATCACTC 1413
Db 80 CAACGCATATCGTGGAGGAACCTGCTCTCTGAAGGTGCTTAAAGCGACTACTGTAAACAC 21

Qy 1414 CGGCAATGGAACCTGCA 1432
Db 20 CAGGACCTTGGCATATCA 2

RESULT 10
US-11-348-413-12794
; Sequence 12794, Application US/11348413
; Publication No. US20060160121A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084100 (AM 101724)
; CURRENT APPLICATION NUMBER: US/11/348,413
; CURRENT FILING DATE: 2006-02-07
; PRIOR APPLICATION NUMBER: PCT/US05/035471
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 11/243,445
; PRIOR FILING DATE: 2005-10-05
```

```
; PRIOR APPLICATION NUMBER: US 60/615,573
; PRIOR FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 1276209
; SEQ ID NO 12794
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(321)
; OTHER INFORMATION: WANO1UOV8; ORF: 00000000011100; Cluster contains WANO1OSHMH
; OTHER INFORMATION: :SE1861:urea:urease gamma subunit:ATCC12228.NC_004461.1
US-11-348-413-12794

Query Match      3.2%; Score 92.6; DB 9; Length 321;
Best Local Similarity 58.7%; Pred. No. 3.5e-12;
Matches 176; Conservative 1; Mismatches 120; Indels 3; Gaps 1;

Qy 207 TGAAGTCTACACCCAAAGAGCAGCAAGAAAAGTTCTTGTATATATATGCGGGGGAAGTGGCTA 266
Db 20 TGCACCTTTACACACAGCTGAACCAAGACAAATTGATGATAGTTGTAGCTGCTGAGGTTGCAC 79

Qy 267 GAAAGCGCAAAAGCAGAGGCGCTTAAAGCTCAACCAACCGGAAGCCATTGCTTACATTAGTG 326
Db 80 GTGCTAGAAAAGCAAGAGACTTAAACTTAATCATCTGAAGCACTTGCTTTAATCAGTG 139

Qy 327 CCCATATTATGGACGAAGCGCGCCGTGGAAAAAACCCTTGCCAGCTTATGGAAGAGT 386
Db 140 ATGATATTATTAGAAGCGCGCGTGATGG--TAAACCGTAGCTGAACIVATGAGCTATG 196

Qy 387 GCATGCACCTTTTGA AAAAAGATGAAGTAATGCCCGGGGTGGTAAATATATGTTCCCGATC 446
Db 197 GAAAAACAATTTTAAACGAGGAAGATGTCATGGATGGCGTAGCTAAACATGATTACAGAAC 256

Qy 447 TAGGTGTAGAGCCACCTTTCTCTGATGGTACGAAACTTGTAACTGTGAATTGGCCCATCG 506
Db 257 TTGAATTTGAAGCAACTTTTCCAGATGGTACTAAGTTAATAACAGTCCATCACCCCAATCG 316

RESULT 11
US-10-471-571A-3255
; Sequence 3255, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: Seqwinn9, version 1.03
; SEQ ID NO 3255
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-3255

Query Match      3.1%; Score 89.8; DB 6; Length 300;
Best Local Similarity 58.3%; Pred. No. 1.6e-11;
Matches 175; Conservative 0; Mismatches 122; Indels 3; Gaps 1;

Qy 207 TGAAGTCTACACCCAAAGAGCAGCAAGAAAAGTTCTTGTATATATATGCGGGGGAAGTGGCTA 266
Db 2 TGCATTTTACACAGAGAGCAAGACAAATTAATGATTGTAGTGGCGGGGGAAGTTCAC 61

Qy 267 GAAAGCGCAAAAGCAGAGGCGCTTAAAGCTCAACCAACCGGAAGCCATTGCTTACATTAGTG 326
Db 62 GTGCTGCTAAGCAGCGTGTGTTGAACATAAATCATCTCTGAGGCAATTAGCTTTAATCAGCG 121

Qy 327 CCCATATTATGGACGAAGCGCGCCGTGGAAAAAACCCTTGCCAGCTTATGGAAGAGT 386
```



```
Db 122 ATGAATTATTAGAGGTGCACCGATGG---TAAGACCGTTGCAGAGTTAATGAGTTATG 178
Qy 387 GCATGCACTTTTGAAGAAAGATGAATATGCCCGGGTGGGTAATATGTTTCCCGATC 446
Db 179 GTACAGAAATCTAAACAAAGAGATGTCATGGATGGTGTGCAACACATGATTACAGATA 238
Qy 447 TAGGTGTAGAGCCACCTTTCTCATGTGTACGAACTTGTAACTGTGAAATGGCCCATCG 506
Db 239 TCGAAATCGAGGCTACGTTCCCGATGGTACTAAGTTAATCAGATACATCACCCCTATG 298

RESULT 12
US-11-348-413-11115
; Sequence 11115, Application US/11348413
; Publication No. US2006016012AI
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084100 (AM 101724)
; CURRENT APPLICATION NUMBER: US/11/348,413
; CURRENT FILING DATE: 2006-02-07
; PRIOR APPLICATION NUMBER: PCT/US05/035471
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 11/243,445
; PRIOR FILING DATE: 2005-10-05
; NUMBER OF SEQ ID NOS: 1276209
; SEQ ID NO 11115
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(402)
; OTHER INFORMATION: WAN01UFWI; ORF: 00000000011110; Cluster contains WAN01OSH
; OTHER INFORMATION: :SE1862:ureb beta subunit:ATCC12228_NC_004461.1
US-11-348-413-11115

Query Match 3.0%; Score 85.2; DB 9; Length 402;
Best Local Similarity 52.8%; Pred. No. 2.2e-10;
Matches 201; Conservative 2; Mismatches 175; Indels 3; Gaps 1;

Qy 543 TTGGTTCGATAAAGCATCGAGCTCAATGCGAGCAAGAGTAACCGAATCTGAGTTA 602
Db 20 TTGTTAAAAATACTGAAATAGAGTCAATAAACATCATCCGAAACGGTGATTGAAGTRA 79
Qy 603 CTAATGAAGGCGCTAAATCCTTTCATGTGGTAGCCATTTCCACTTCTTTTGAAGCTAACA 662
Db 80 AATAACRGGGATAGACCTATACAGTAGGTTCCACTTTTTCGAAACAATA 139
Qy 663 AGGCACATAAATTCGATCGTGAAGAAAGCCTATGCGAAACGCTAGATATTCCTCTGCGCA 722
Db 140 ARGCAATTAGATTTGATCGTGAGAAACATATGTAACATTTGGATATCTCTCGAGAG 199
Qy 723 ACAGCTACGATTTGGGGAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGCGCA 782
Db 200 CTGCAGTGAGATTTGAACCTTGGAGATGAAAAAAGTACAACTTGTGCAATATTTCTGGAC 259
Qy 783 GTAAAAAGCTGATGGCATGACGCGGCTTGTGAAT---AACATCGCGATGACGCAATA 839
Db 260 GACGTAAAAATTTATGGATTCGTTGGTGTAGTCGATGCGGATATTTGACGAAACGCGTAT 319
Qy 840 AACATAAAGCGCTTGACAGGCGGAAATCTCACCGGATTTATCAAGTAAGGAGACTCCCATG 899
Db 320 TCGTTCGAATGATTCAAATCAAAACGCCCGGTTAAAAACGATGACGAGGAGACAATG 379
Qy 900 AAAATGAAAAAACAAGATAT 920
Db 380 CGAATAAAAAAGGTGTAAT 400
```

```
RESULT 13
US-10-471-571A-3277
; Sequence 3277, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 3277
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-3277

Query Match 2.8%; Score 80.2; DB 6; Length 408;
Best Local Similarity 50.9%; Pred. No. 3.3e-09;
Matches 190; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

Qy 553 TAAAGACATCGAGCTCAATGCGAGGCAAGAAAGTAACCGAATCTGAGTTACTAATGAAG 612
Db 30 TACAGAGGTTGAAATTAATAACCATCATCTCTGAAACAGTTATCGAAGTTGMAAATACAGG 89
Qy 613 GCCTAAATCCTTTCATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAACAGGCACATAA 672
Db 90 AGATCGACCAATTCAGTGGGGCTCACATTTTCATTTTATGAAGCAAAATGCGAGCATTAGA 149
Qy 673 ATTGCATGCTGAAAGAGCTATGCGAAAGCCTAGATATTCCTCTGCGCAACACGCTACG 732
Db 150 TTTCGACGTGAAATGGCATATGAGAAACATTTAGATATTCAGCTGGAGAGCTGTTCG 209
Qy 733 CATTTGGGCGAGCAACAAACCCGAAAGTGCAGTTGATTCCTCTTTGGTGGCAGTAAAAAAGT 792
Db 210 ATTGAACTGGGATAAAGAAAGATTCATTTAGTATGCTGGCAACGTTAAAT 269
Qy 793 GATTGGCATGAACGGGCTTGTGAATAACATCGCGATGAACCGCCATAAAACATAAAGCGCT 852
Db 270 TTTTGGTTTTCGTGTATGGTCAATGGTCTATCGATGAGTCACGCTGTCTATCGCCCAAC 329
Qy 853 TGACAGCGGAAATCTACGGATTTATCAAGTAAGGAGACTCCCATGAAATGAAAAAAC 912
Db 330 TGATGAAATGATGAATATGCAAGGTGTATTTCGAGATAACCGGTGCTGAAAAACGTGAATA 389
Qy 913 AAGAAATATGTAAA 925
Db 390 AAAGAGGAGAAA 402

RESULT 14
US-11-348-413-11114/c
; Sequence 11114, Application US/11348413
; Publication No. US2006016012AI
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084100 (AM 101724)
; CURRENT APPLICATION NUMBER: US/11/348,413
; CURRENT FILING DATE: 2006-02-07
; PRIOR APPLICATION NUMBER: PCT/US05/035471
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 11/243,445
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 60/615,573
```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2006, 01:16:43 ; Search time 12167 Seconds
(without alignments)

13250.221 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2880.6

Sequence: 1 rgragattttccaractt.....aaaaaagtagaacacagg 2883

Scoring table:

IDENTITY_NUC

Gapop 3.0 , Gapext 3.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est3:*

3: gb_est4:*

4: gb_est5:*

5: gb_est6:*

6: gb_estc:*

7: gb_est2:*

8: gb_est7:*

9: gb_est8:*

10: gb_est9:*

11: gb_gss1:*

12: gb_gss2:*

13: gb_gss3:*

14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c					
1	364.8	12.7	956	14	DU750758 ASNF1644.
2	346.2	12.0	936	14	DU760951 ASNG3437.
3	335.2	11.6	942	14	DU762934 ASNG641.g
4	333.2	11.6	958	10	DR637438 EST102806
5	330.2	11.5	1509	6	BR816461 Arabidops
6	330.2	11.5	946	10	DR637093 EST102771
7	320	11.1	809	10	DV133265 CV03067A1
8	312.2	10.8	925	10	DR633422 EST102404
9	309	10.7	872	8	CO122599 GR_EB04F
10	306.8	10.7	889	10	DV759139 PchrSEQ10
11	302.2	10.5	748	9	CK676965 ydd29R04.
12	300.2	10.4	889	10	DR637386 EST102801
13	296.4	10.3	796	10	DR633939 EST102456
14	293	10.2	849	10	DR635599 EST102622
15	284	9.9	784	11	AZ935182 BJ_Ba000
16	279.6	9.7	766	8	CV195867 CGF10344
17	276.4	9.6	845	10	DV135585 CV03111B2
18	272.4	9.5	784	10	DV758961 PchrSEQ87
19	264.6	9.2	713	5	CK757406 atr02-4ms

c	20	264.4	9.2	986	5	CF826586	EST703968
c	21	262	9.1	946	9	DN809412	73954613
c	22	261.2	9.1	684	8	DN907928	030109ABL
c	23	261	9.1	921	5	CF714848	CCAF343TO
c	24	260.8	9.1	945	9	DN808832	76945303
c	25	259.8	9.0	863	5	CD375065	PTMM00226
c	26	253.2	8.8	905	4	CA264652	SCOGLB204
c	27	252.4	8.8	796	5	CF708414	CCABR70TR
c	28	250.8	8.7	689	10	DR440006	EST1149_09
c	29	249.4	8.7	757	3	BU027777	QHG7L07.Y
c	30	248.6	8.6	781	10	DR916924	EST110846
c	31	248.4	8.6	739	10	DV152796	CV03090A1
c	32	247.8	8.6	661	8	DN908212	030109ABL
c	33	243.2	8.4	789	4	CA236757	SCEQFI505
c	34	240.4	8.3	692	8	CV902222	PD019F1 m
c	35	239.8	8.3	937	9	DN810553	76954039
c	36	239.4	8.3	990	8	CO027797	EST806181
c	37	239	8.3	643	14	AG940016	Drosophil
c	38	237.2	8.2	634	2	BG887260	EST513111
c	39	236.8	8.2	1275	12	BZ572656	msb2_2738
c	40	235.8	8.2	660	11	AZ935008	BJ_Ba000
c	41	235.4	8.2	829	5	CF689903	CCADS40TF
c	42	234.8	8.2	636	4	CA851837	D18B02.D1
c	43	234.2	8.1	626	8	CN910029	030124ABL
c	44	231.8	8.0	614	11	AZ935371	BJ_Ba000
c	45	227.4	7.9	747	5	CF690511	CCAAZ91TR

ALIGNMENTS

RESULT 1
DU750758/c
LOCUS
DEFINITION ASNF1644.b2 HF130_10-06-02 uncultured marine microorganism
HF130_10-06-02 genomic clone HF0130_019D02, genomic survey
sequence.

ACCESSION DU750758
VERSION DU750758.1 GI:85760594
KEYWORDS GSS.
SOURCE uncultured marine microorganism HF130_10-06-02
ORGANISM uncultured marine microorganism HF130_10-06-02
REFERENCE 1 (bases 1 to 956)
AUTHORS Delong, E.P., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J.,
Frigaard, N.U., Martinez, A., Sullivan, M., Edwards, R., Chisholm, S.W.
and Karl, D.M.

TITLE Comparative genomics reveals ecological trends in stratified
microbial communities in the ocean's interior
JOURNAL Science (2006) In press
COMMENT Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardson
and Edward DeLong

US DOE Joint Genome Institute
2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
Tel: 617-253-5271
Fax: 617-253-2679
Email: PRRichardson@lbl.gov; delong@mit.edu

North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid
DNA library prepared from marine picoplankton in the less than 1.6
um, greater than 0.22 um fraction. Sample Date: 10/6/2002
Coordinates: 22.45 N, 158 W Depth 130 m Temperature: 22.19 C
Salinity: 35.31 psu Oxygen: 204.9 umol/kg
Class: fosmid ends.

FEATURES
source
1. 956
/organism="uncultured marine microorganism HF130_10-06-02"
/mol_type="genomic DNA"
/db_xref="taxon:361146"
/clones="HF0130_019D02"
/cell_type="marine picoplankton, less than 1.8 um, greater
than 0.22 um fraction"

ORIGIN		Query Match		Best Local Similarity		Matches 588; Conservative		Indels		Gaps	
		12.7%; Score 364.8; DB 14; Length 956;		61.8%; Pred. No. 2.6e-75;		0; Mismatches 362;		1;		1;	
QY	1250	AGCACTAGCAGGGAAGGTATGATTATTACCGCTGGGGGAATCGATTTCACACACCCCACTT	1309								
DB	956	ACCATTCGAGGAGGAAGAGCATATTATCTGAGGAGCTATTGATAGTCACATTCACCTT	897								
QY	1310	CTTTTCTCCACAACAAATTCCTACCGCTCTAGCCAATGGCGTTTACACACCATGTTGGAGG	1369								
DB	896	AATCTGTCTTCAACAAATAGAAACTGCTTTAGCTAGTGGGGTTTACCACAATGCTCGAGGG	837								
QY	1370	CGGCACAGGTCCTGTAGATGGCAGGAATGGCACTACTATCATCTCCGGGCAAAATGGAACTT	1429								
DB	836	AGGACAGGCGCCAGCAACAGGTACTAATGCAACGACATGTACACCTGGAGCATTTCCACAT	777								
QY	1430	GCACCGCATGTTCGGCGCACAGAAGAGTATTCATGAATATGCGGCTTTTGGGCAAGG	1489								
DB	776	CTCAGAAATGCTGCANATCTCGAGAGGATTTCTGTTAATTGGGATCTTTGGAAAGG	717								
QY	1490	CAATPAGCTCTAGCAAAAAACAACCTGTGTAGAACAAAGTAGAAGCGGGCGCATTTGGTTTTAA	1549								
DB	716	CAATGCAACTAAACAAGCAGCATTTAGAAGAACAAAGTAAGAGCAGGTCTTTGGGTTAAA	657								
QY	1550	ATTGCATGAAGACTGGGGCAACAACAAGTGCGATCGATCACTGCTTGAGCGTGGCAGA	1609								
DB	656	ACTCCATGAAGACTGGGGAAACAACACGGGCTGTATTGATTTCTGCTTAAGTGTGCAGA	597								
QY	1610	TGAATACGATCTGCAAGTTTGATCTCCACACCGATACAGTCAANTGAGCGAGTATGTAGA	1669								
DB	596	TCAACTAGACGTACAGTTTGATTTATTCATACAGATACCCTAAATGAAGCTGGTTTGTGTA	537								
QY	1670	TGACACCTTAAATGCAATGAACGGCGCGGCATCCATGCGCTACACATTCAGGGAGCGGG	1729								
DB	536	AGATACAAATTAAAGCANTAGAGGAGAAACAATTCATACCTTCCATACAGAGGGCTGG	477								
QY	1730	TGGAGGACACTCACCTGATGTTATCACATGGCAGGCGAGCTCAATATTCTACCTCCTC	1789								
DB	476	AGGTGGTCAAGCTCCCGACATATAAAAAATTTGTGGAGAAATCAAAATGTGATTTCCACGAG	417								
QY	1790	CACCACCCCACTATTCCTTATACCATTAATACGGTTGCGAGACACTTACAGCATGCTCAT	1849								
DB	416	TACAAATCCAACTAGGCTTTTCACTCTAAATATCTTTTGAAGAGCATTTAGACATGTTGAT	357								
QY	1850	GACATGCCACCACTAGACAAACGATCCCGAGGATTTTCAATTTTCTCAAAGCCGTAT	1909								
DB	356	GGTTTGTATCATTTTAGATCCCAAAATTCAGAGGATGTTGCATTTGCTGAGTCAGAAAT	297								
QY	1910	CCGCCCCGGCTCTATCCGGCTGAAGATGTCTCCATGATATGGGTGTGATCGCGATGAC	1969								
DB	296	ACGTCGTGAAACTATTCTGCTGAGGACATCTCCACGACTTAGAGCGCTTTTCTATTAT	237								
QY	1970	AAGCTCGGATTCGAACGAATGGGGCGTGCAGCGGAAGTATTCCTCGAACTTGGCGAGC	2029								
DB	236	TGCTAGTGACTCCACAGCTATGGGTAGAGTTGGGGAGGTGATTAAGCCGGACTTTTCAAAC	177								
QY	2030	TGCGGATAAGAAATAAAAAAGAAATTTGTGTAAAGCTTCCTGGAAGATGGCAAGATAACGATAA	2089								
DB	176	TGCTCATAAATGAAGTTCAAAGAGAGCCCTTACCTTGAGGATTAATCAAAGGAATGATAA	117								
QY	2090	TTTCCGCAATTAAGCGCTACATCTCCAAATAACATPATCAACCCCGCTTTGACCCACGCGGT	2149								

Db	116	TCATCGTCTGAAAGATATATCTCAAAGGTCACCTAATTAATCTCGCGANTAGCTCATGGAAT	57
Qy	2150	GAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTGGT	2200
Db	56	CATGCTCATGTTGGTTCGGTAGAG-TTGAAAACTAGCAGACTTAGTGAT	7
RESULT 2			
LOCUS	DU760951		
DEFINITION	ASNG3437.g2 HF200_10-06-02 uncultured marine microorganism HF200_10-06-02 genomic clone HF0200_094E10, genomic survey sequence.		
ACCESSION	DU760951		
KEYWORDS	DU760951.1	GI:85770787	
SOURCE	GSS.		
ORGANISM	uncultured marine microorganism HF200_10-06-02 uncultured marine microorganism HF200_10-06-02 unclassified sequences; environmental samples.		
REFERENCE	1 (bases 1 to 936)		
AUTHORS	DeLong,E.F., Preston,C.M., Mincer,T., Rich,V., Hallam,S.J., Frigaard,N.U., Martinez,A., Sullivan,M., Edwards,R., Chisholm,S.W. and Karl,D.M.		
TITLE	Comparative genomics reveals ecological trends in stratified microbial communities in the ocean's interior		
JOURNAL	Science (2006) In press		
COMMENT	Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus, Kerrie Barry, Tijana Glavinadelirio, David Bruce, Paul Richardson and Edward DeLong US DOE Joint Genome Institute US DOE Joint Genome Institute 2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA Tel: 617-253-5271 Fax: 617-253-2679 Email: PWRichardson@lbl.gov; delong@mit.edu		
FEATURES	North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid DNA library prepared from marine picoplankton in the less than 1.6 um, greater than 0.22 um fraction. Sample Date: 10/6/2002 Coordinates: 22.45 N, 158 W Depth 200 m Temperature: 18.53 C Salinity: 35.04 psu Oxygen: 198.8 umol/kg Class: fosmid ends.		
source	Location/Qualifiers		
	1..936		
	/organism="uncultured marine microorganism HF200_10-06-02"		
	/mol type="genomic DNA"		
	/db xref="taxon:361147"		
	/clone="HF0200_094E10"		
	/cell type="marine picoplankton, less than 1.8 um, greater than 0.22 um fraction"		
	/clone lib="HF200_10-06-02"		
	/notes="Vector: pCG1FOS; North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid DNA library prepared from marine picoplankton in the less than 1.6 um, greater than 0.22 um fraction. Picoplankton collected at 200 m depth on 10/6/2002, Coordinates: 22.45 N, 158 W. Sample Date: 10/6/2002 Coordinates: 22.45 N, 158 W Depth 200 m Temperature: 18.53 C Salinity: 35.04 psu Oxygen: 198.8 umol/kg"		
ORIGIN			
Query Match	12.0%; Score 346.2; DB 14; Length 936;		
Best Local Similarity	60.8%; Pred. No. 6.9e-71;		
Matches	564; Conservative	0; Mismatches 363; Indels	0; Gaps 0;
Qy	1665	GTGATGACACCCCTAAATGCAATGAACGGGGCGGCATCCATCGCTACCACTAGGGA	1724
Db	2	GTGAGAATACCGTCGCGCGATGAACGCGGTCATCCACGCTTCCATACTGAGGGG	61
Qy	1725	GCGGGTGGAGACACTCACTCGATGTTATCACCATGCGAGCGAGCTCAATATTCTACCC	1784
Db	62	GCCGGTGGCGGCACGCGCGGACATCATCAAGATCTGCGGCGAAGCACTAGTCTCCG	121
Qy	1785	TCCTCCACCAACCCCCCACTATTCCCTTATACCATTAATATACGGTTGCGAAGCACTTAGACATG	1844

```
Db 122 TCATCGACCAACCGGCGCCCTACACGCTGAACAGCTTGAAGCATCTCGACATG 181
Qy 1845 CTATGACATGCCACACCTAGACAAAGCATCGCGAGGATTTACAAATTTCTTCAAGC 1904
Db 182 CTATGCTCTGCCATCACTCGACAACTCAATACCGGAAGATGTGGCTTTGGCGGAAGC 241
Qy 1905 CGTATCGCGCCCGCTCTATCGCGGCTGAGATGCTCCATGATATGGGTGATCGCG 1964
Db 242 CGATCAGGCGGCAACATTCGCGGAGAGACATCTTTCATGACATGGGGGCTTTTCG 301
Qy 1965 ATGACAAGCTCGGATTCGCAAGCAATGGGCGTGCAGCGCAAGTGAATCTTCGAACCTGG 2024
Db 302 ATCATCGGCTCGACACCGAGGCATGGCGGTTCGCGAGGTTCATCATCGGACCTGG 361
Qy 2025 CAGACTCGGATAGAAATAAAGAAATTTGGTAAGCTTCTGGAAGTGGCAAGATAAC 2084
Db 362 CAGACAGCGCAAGAATGAAGTCCAGCGCGGAAGGCTGCTGAGGAACCGGCGCAAC 421
Qy 2085 GATAATTTCCGATTAAGCGCTATCTCCAAATACATCACTATCAACCCGCTTTGACCCAC 2144
Db 422 GACAATTTGCGGGTCAAGCGCTATGTCGGAATATACCACTCAACCCGCGCATCGCCAT 481
Qy 2145 GCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTGTGTGG 2204
Db 482 GGATATCAGCGCATATTCGGGTCTGATCGCGTGGGAAGCGGCGGACCTTGTGCTCTGG 541
Qy 2205 AATCTGCGCTTTTGGCGTAAACCCAAATCGTGATCAAGGCGGTATGGTGTCTTC 2264
Db 542 AATCGCGCTTCTTGGCGTGAAGCGGAGATGTCATCATCGGCGGCGAGCATGCTCTGC 601
Qy 2265 TCTGAATGGCGGATTTAAGCGCTGTGTGCCACTCCCAACCGGTTTATACCGGAA 2324
Db 602 GCGCAGATGGCGGATCTAAGCGCTCAATCCGACACCGCAGCGCTGTCTACACCGCGCG 661
Qy 2325 ATGTTTGGGATCAAGCGAGGGAATTTGACACAGCATCACTTTTGTTCAAAGTC 2384
Db 662 ATGTTGCGCGGCTTTCGCGAGTGGTGGAGAAATTCGGCTGTCACTTCGTGAGCGCAAG 721
Qy 2385 GCCTATGAAATGGCGTGAAGAAAGCTGGCTTAGAGCGGCAAGTTCTACCGGTCAAA 2444
Db 722 GCGCAGATGCGCGATCGCGGACACGCTCGGCTTCCAAAGGCGACCTCGCGTGGCG 781
Qy 2445 AACTGCGGTAACATCAACGAAGAAGCTTCAAGTTCAACGACAAAGCGCAAAATCAAC 2504
Db 782 AAGACGCTACGATCTCCAGGCGGACATGTTGCATAACGATTACTGCCCCGACATCGAG 841
Qy 2505 GTCGATCCGAAACCTTCGAGGCTTTGTAGATGGCAACTCTGCACTCTAAACCCACC 2564
Db 842 GTCGACCCGAAACCTATGAGGTTTCGGCTGATGGCGAGCTTCTGACCTCGAGCGCGCG 901
Qy 2565 TCGCAAGTGCCTTAGCCCGCGCTAC 2591
Db 902 GAAGAACTGCGGATGCGGAGCGCTAC 928
```

```
RESULT 3
LOCUS DU762934
DEFINITION 942 bp DNA linear GSS 27-JAN-2006
ASN641.g2 HF200_10-06-02 uncultured marine microorganism
HF200_10-06-02 genomic clone HF0200_064A09, genomic survey
sequence.
ACCESSION DU762934
VERSION DU762934.1 GI:85772770
KEYWORDS GSS.
SOURCE uncultured marine microorganism HF200_10-06-02
ORGANISM uncultured marine microorganism HF200_10-06-02
unclassified sequences; environmental samples.
REFERENCE 1 (bases 1 to 942)
AUTHORS DeLong,E.F., Preston,C.M., Mincer,T., Rich,V., Hallam,S.J.,
Frigaard,N.U., Martinez,A., Sullivan,M., Edwards,R., Chisholm,S.W.
and Karl,D.M.
TITLE Comparative genomics reveals ecological trends in stratified
```

microbial communities in the ocean's interior
Science (2006) In press
Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
Kerrie Barry, Tjiana Glavinadelrio, David Bruce, Paul Richardson
and Edward DeLong
US DOE Joint Genome Institute
2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
Tel: 617-253-5271
Fax: 617-253-2679
Email: PWRichardson@lbl.gov; delong@mit.edu
North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid
DNA library prepared from marine picoplankton in the less than 1.6
um, greater than 0.22 um fraction. Sample Date: 10/6/2002
Coordinates: 22.45 N, 158 W Depth 200 m Temperature: 18.53 C
Salinity: 35.04 psu Oxygen: 198.8 umol/kg
Class: fosmid ends.

FEATURES
source
1..942
/organism="uncultured marine microorganism HF200_10-06-02"
/mol_type="genomic DNA"
/db_xref="taxon:361147"
/clone="HF0200_064A09"
/cell_type="marine picoplankton, less than 1.8 um, greater
than 0.22 um fraction"
/clone_lib="HF200_10-06-02"
/note="Vector: pCC1FOS; North Pacific Subtropical Gyre
(Hawaii) picoplankton genomic fosmid DNA library prepared
from marine picoplankton in the less than 1.6 um, greater
than 0.22 um fraction. Picoplankton collected at 200 m
depth on 10/6/2002, Coordinates: 22.45 N, 158 W. Sample
Date: 10/6/2002 Coordinates: 22.45 N, 158 W Depth 200 m
Temperature: 18.53 C Salinity: 35.04 psu Oxygen: 198.8
umol/kg"

ORIGIN

Query Match 11.6%; Score 335.2; DB 14; Length 942;
Best Local Similarity 62.6%; Pred. No. 2.8e-68;
Matches 538; Conservative 0; Mismatches 318; Indels 3; Gaps 1;
Qy 1308 TTCCTTTCTCCACAACTCCCTACCGCTCTAGCCAACTGCTTACACCACTGTTTGA 1367
Db 1 TTTATCTCGCCCAATCTTATTTCTGAAGCTTCTTATTCGGCATCACCACTCATCTGC 60
Qy 1368 GCGCGCACAGTCTCTAGATGGCAAGATGCGACTTATCTACCTCCGGGCAAAATGGAAC 1427
Db 61 GGTGGCAGGGGCGGCGACTGGCACCAACGCGACCACTGCACTCCGCGTGGGAAC 120
Qy 1428 TTGCACCGCATGTTGGCGCGCAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAA 1487
Db 121 CTCGCAAAATGTACGAGGCGGTGGAGCGTTCCCGCTCAACTTCGGCTTTTGGGCAAG 180
Qy 1488 GGCATAGCTCTAGCAAAACAACTTGTAGAACAGTAGAGCGGCGCGATGGTTT 1547
Db 181 GGCACAGCTCGCTTCCAAATGCCCTTGGGGAACAGGTGGAAGCGGAGCGATGGCCCTG 240
Qy 1548 AAATTGATGAAGACTGGGGCACAAACCAAGTGGATCGATCACTGTGTAGCGGTGGCA 1607
Db 241 AATTGACAGAGACTGGGGCACAGACTCCGGGCGCATGTACTGTCTCGCGTGGCGG 300
Qy 1608 GATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGCAGGTTATGTA 1667
Db 301 GAGGAGTACGACGTGCGAGGTGCGCATCCACACCGACCGCTCAACGAATTCGGCTTGTG 360
Qy 1668 GATGACACCTTAATGCAATGAACGGGCGCGCATTCATGCTTACCATGAGGAGCG 1727
Db 361 GAGGACACGCGGCGCGGTTTCAAGGAGACGACCATCCACATCCACACATGAGGCGCA 420
Qy 1728 GGTGGAGGACACTCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTCTACCTCC 1787
Db 421 GCGCGCGACACGCGCCCGGACATCATCAAGCTCTCGCGGCGAGGCCACGATGCTGCTCC 480
Qy 1788 TCCACCACCCCACTATTTCCTATATACCAATTAATACGGTTGAGAGAACACTTAGACATGCTC 1847

```
Db 481 TCACCAATCCGACGATGCGGTTACCGCAACACGATGACGAGCATCTGGACATGCTG 540
QY 1848 ATGACATGCCACCACTAGACAAACGATCCGCGAGGATTTACAAATTTCTCAAAGCGGT 1907
Db 541 ATGGTCTGCACCACTCTCGCCCGGATTCGCGAGGACGTGCGCTTTGCGAATCGGC 600
QY 1908 ATCCGCCCGCGCTTATCGCGGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATG 1967
Db 601 ATCCGCCCGGAGACCAATTCGCGCGGAGACATCTGCACGACCTCGCGGCGTTTCAAGATT 660
QY 1968 ACAAGCTCGAATTCGCAAGCAATGGGCGGTGCGAGGCAAGTATCTTCGAACTTGGCAG 2027
Db 661 CTGCGCTCGAATTCGCAAGCGATGGCGCGCATCGCGCAAGTGAATTCGCGCACCTGGCAG 720
QY 2028 ACTGCGGATAAGAAATAAAAGAAATTTGGTAAGTCTCTGAAGATGCAAAAGATAACGAT 2087
Db 721 ACCGCCGACGAAGTAGAGCCCGACGCGGCTGCTCCGGCGACGCG--GGCAGGAC 777
QY 2088 AATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTTGACCCAGGC 2147
Db 778 AACCTGCGCATCGCTGCTAGCTCCGCAATACACCATCAACCCCGCATCGCCCAACGCGC 837
QY 2148 GTGAGCGAGTATATCGGCT 2166
Db 838 ATAAACAAGGTCGTTGGCT 856

RESULT 4
DR637438
LOCUS
DEFINITION
  DR637438 958 bp mRNA linear EST 11-JUL-2005
  EST1028063 FvM Gibberella moniliformis cDNA clone FVMAW64, mRNA
  sequence.
ACCESSION
  DR637438
VERSION
  DR637438.1 GI:70712272
KEYWORDS
  EST.
SOURCE
  Gibberella moniliformis
  Gibberella moniliformis
  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
  Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE
  1 (bases 1 to 958)
  Brown,D.W., Cheung,F., Proctor,R.H., Butchko,A.E., Zheng,L, Lee,Y.,
  Utterback,T., Smith,S., Feldblyum,T., Glenn,A.E., Plattner,R.D.,
  Kendra,D.F., Town,C.D. and Whitelaw,C.A.
  Comparative analysis of 87,000 expressed sequence tags from the
  fumonisin-producing fungus Fusarium verticillioides
  Fungal Genet. Biol. 42 (10), 848-861 (2005)
  16099185
  Contact: Brown, D.W.
  USDA/ARS/NCAUR
  1815 N. University St, Peoria, IL 61604, USA
  Tel: 309 681 6230
  Fax: 309 681 6689
  Email: brown@ncaur.usda.gov
  TIGR sequence name: FVMAW64TH
  Seq primer: AAT TAA CCC TCA AAG GG.
  Location/Qualifiers
    1..958
    /organism="Gibberella moniliformis"
    /mol_type="mRNA"
    /strain="m3125"
    /db_xref="taxon:117187"
    /clone_lib="FvM"
    /tissue_type="mycelia"
    /note="Vector: pBlueScript II SK(+) XR; Site 1: EcoRI;
    Site 2: XhoI; anamorph: Fusarium verticillioides. Library
    FvM was prepared from pooled RNA obtained from a 48-hour
    and a 72 hour, liquid GYM culture from strain M-3125.
    Cultures were vacuum filtered and the mycelial mats were
    frozen in liquid nitrogen, ground to a powder, and then
    added to TRIzol Reagent (Invitrogen, Carlsbad CA) at
```

```
approximately 1 g mycelia per 10 ml TRIzol. The cDNA was
directionally ligated into the pBlueScript II SK(+) XR
vector (cDNA Synthesis Kit; Stratagene)."
```

ORIGIN

Query Match 11.6%; Score 333.2; DB 10; Length 958;
Best Local Similarity 60.6%; Pred. No. 8.5e-68;
Matches 572; Conservative 0; Mismatches 368; Indels 4; Gaps 2;

QY 1209 GATGCGTAAGCCCTCATATGTCGTGGTGTGGGCACAGAAAGCACTAGCAGGGGAAGGT 1268
Db 2 GACGCGTAAACAGAGGCGATGTCGCGAAGTCACGCGATGTCGTACAGGTGAAGGA 61
QY 1269 ATGATTAATTAACCGTGGGGGAATCGAATTCACACCCCACTTCTCTTCTCCACAACAATTC 1328
Db 62 AAGATTGTGACCGCGGCGCTATCGACAGCATATTCACCTTATCTGCTCCCTCAGCAAGTA 121
QY 1329 CTTACCGCTCTAGCCAAATGGCGTTTACCAACCATGTTTGGAGGCGGCACAGTCTCTAGAT 1388
Db 122 CCGAAGCTCTTGCATCTGCTGTAAACCAACCATGCTTGGCGGTGGTACCGGCCCAAGTGT 181
QY 1389 GGCACGAATCGACTTACTATCACTCCGGGCAATGGAACCTTCACCGCATGTTTGGCGCA 1448
Db 182 GGAACGAACGCAACTTCTGACGCTGTGCTCATTAATGCTCAATGTTTGCAGGCG 241
QY 1449 GCAGAAGAGTATTTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAA 1508
Db 242 TCGCATCAGCTTCTTATCAATATTGTTACTTGGCAAGGTAATGATAGTTCTCTCTGAG 301
QY 1509 CAATCTGTAGAACAAAGTAGAGCGGGCGGATGTTTAAATTTGCATGAAGACTGGGGC 1568
Db 302 GGTCTGCGCGATCAGGTCAATGCTGCTGTGGCTCAAGCTTCATGAGGACTGGGT 361
QY 1569 ACAACACCAAGTGGCATCGATCAGTCTTACGCGTGGCAGATGAATACCATGTGCAAGTT 1628
Db 362 TGCATCTCTGCTATTGACGCTTGTCTCAGTGTCTGTGATGAATTCGATATTCATGT 421
QY 1629 TGTATCCACACCGATACAGTCAATGAGGCGGTTATGTAGATGACACCCCTAAATGCAATG 1688
Db 422 CTATTACACTGACACGCTTAACGAGTGTGCTTGTGCGAATCTACGATCGCTGCTTTC 481
QY 1689 AACGGGCGCGCATCCATCGCTTACCAATTAGAGGAGCGGTGGAGGACACTCACCTGAT 1748
Db 482 AAGAACCCGCAAAATTCATCTTATCACAGAGGGTGCAGGAGGTGGCCATGCTCCGGAT 541
QY 1749 GTTATCACATGGGAGGCGAGTCAATATTTCTACCTCTCCACACCCGCCATTTCCC 1808
Db 542 ATCATCTCCGGTGTAGAGCATCAAAATGTTCTGCCATCATCGACCAACCTTCAAGACCA 601
QY 1809 TATACCATTAATACGGTTGCGAAACACTTTAGACATGCTCATGACATGCCACCATCTAGAC 1868
Db 602 TTCACACGCAATACTCTCGATGAGCATCTCGATATGCTTATGCTCGCATCACTTGTTC 661
QY 1869 AACCGCTCCGCGAGGATTTACAATTTTCTCAAAGCGGTATCCGCCCGGCTTATCGCG 1928
Db 662 AAGAATAATCCAGAGGATGTAGCCCTTCGCGGAGAGCGGTATTTCGTGCTGAAACCAATTGT 721
QY 1929 CTTGAAGATGCTCTCCATGATATGGTGTGATCGGATGACAGCTCGGATTCGGAAGCA 1988
Db 722 GCTGAGGATGTTATACGACAAAGGCGCTATCAGCATGATGAGTCTCACTCGCAGGCT 781
QY 1989 ATGGGCGTGCAGGCGAAGTGAATTCCTCGAATCTGCGACACTGCGGATGAAGATAAAAAA 2048
Db 782 ATGGCCCGTTGCGGAGAGGTGTTTAAAGAACATGGAATCTGCGCATAAATAAGGTG 841
QY 2049 GAATTTGGTAAGCTTCTCTGAAGAT---GGCAAGATAACGATAATTTCCGCATTAAGCGC 2105
Db 842 CAGAGGGGTGTTGTCGCGAGGATGAAGGSCACAGGGGCTGATAATGCGCG-TGTAAACGC 900
QY 2106 TACATCTCAAATACACTATCAACCCCGCTTTGACCCAGCGGT 2149
Db 901 TATGTGACGAAGTATACTATTAAACCCAGCTATTGCTCAGGCGTT 944

RESULT 5
 CNSOABR8 1509 bp mRNA linear HTC 06-FEB-2004
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 DEFINITION GSLTPCH42D08 of Hormone Treated Callus of strain col-0 of
 Arabidopsis thaliana (thale cress).
 ACCESSION BX816461.1 GI:42471655
 VERSION HTC; GSLT cDNA.
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 1509)
 AUTHORS Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
 Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
 Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
 TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1509)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URGV-INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.
 FEATURES
 source
 Location/Qualifiers
 1..1509
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="GSLTPCH42D08"
 /tissue_type="Hormone Treated Callus"
 /ecotype="Col-0"
 /plasmid="pCMVSPORT_6"
 1..1509
 /gene="At1g67550"
 gene
 1..1509
 /gene="At1g67550"
 ORIGIN
 Query Match 11.5%; Score 330.2; DB 6; Length 1509;
 Best Local Similarity 59.1%; Pred. No. 5e-67;
 Matches 602; Conservative 0; Mismatches 413; Indels 4; Gaps 4;
 QY 1587 GATCCTGCTTGAGCGTGGCAGATGATACGATGTGCAAGTTTGTATCCACCGGATACA 1646
 DB 1 GACAAATGTTTGGCAGTTGCGAAGATATGATATCCAGGTGAACATTCATCATGACACC 60
 QY 1647 GTCATGAGCAGGTTATGATGATGACACCTTAAGCAATGAACGGCGCGGCATCAT 1706
 DB 61 TTAATGAATCCGGTTTGTGTAGCACATCATCAATGCAATTCGTGGGAGAACAAATACAT 120
 QY 1707 GCCTACACATTCAGGCGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGGC 1766
 DB 121 ACATATCACAGTAAGGTGCTGGTGGGACATGCACCATATCATATAGAGTTGTGGA 180
 QY 1767 GAGCTCAATATTCCTCCTCCACCCACCTATTCCTTATACCATTAATACGGTT 1826

Db 181 GTGAAAATGTACTCCCGTCATCAACCAACCCCAACGCCATATACGAAAATACTGTA 240
 QY 1827 CGAGAACAACCTTAGACATCTCATGATGATGCGCACCACTTAGACAAACGCATCCGCGAGGAT 1886
 Db 241 GATGAACATCTTGACATGCTGATGTTTGGCCATCACCCTTGACAAGAACATCCCGAAGAT 300
 QY 1887 TTACAAATTTTCTCAAAGCGGTATCCGCCCGGCTCTATCGCGCTGA-AGATGTGCTCCA 1945
 Db 301 GTAGCTTTTGTGTAATCAAGAATAAGAGCTGAACAATAGCTGCAGAGAGATATATTGCA 360
 QY 1946 TGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGGTGCAGCGCA 2005
 Db 361 TGATATGGAGCAATCAGCAATATCTCTCCGACTCACAAGCAATGGGTGCGATTGGAGA 420
 QY 2006 AGTGATTCCTTCGAACCTTCGCAGACTGCGGATGAAGATAAAAAAGAAATTTGGTAAGCTTC 2065
 Db 421 GGTAATAAGCAGAACTTGGCAGACAGCTGACAAGATGAAGGCTCAGAGGGGAGCCATTA 480
 QY 2066 TGAAGATGGCAAGATAAGATAAGTAATTTCCGCAATTAAGC-GCTACATCTCAATACACTA 2124
 Db 481 TCCTAACATGGCTGATGATGATAACTCAAGGATCAAGCAGATACATTTGCAAGATGACACTA 540
 QY 2125 TCACCCCGCTTTGACCCCAACCGCGTGAGAGTATATCGCTCTGTGGAAGAGGGCAAGA 2184
 Db 541 TAAACCCAGCTATAGCCAAATGGGTTTGAGACTTAATTCGCTCGGTTGAGTGAAGAGC 600
 QY 2185 TCGCGCACTTCGTGTGTGGAATCC-TGCTTTTTCGGGTAAAAACCCAAAAATCGTGATC 2243
 Db 601 TGGCTGATCTTGTATATATGCGCAGCGCGCTTCTTCGGAGCAAAACCCAGAAATGATAATC 660
 QY 2244 AAAGCGGTATGGTGTCTTCTGAAATGGCGAATTTAAACGGTCTGTGCGCACTCC 2303
 Db 661 AAAGGAGCAATATAGCATGGCTAATATGGCGATGCAAAATGCGAGCATACCCACTCT 720
 QY 2304 CAACGGTATTATACCGCAATGTTTGGGCATCAGCGCAAGCGGCAAAATTTGACACCAGC 2363
 Db 721 GAGCGGTGATATCAAGACCTTATTTGGAGCAATTTGGAAGGCGCGGAAGTGAAGAACTCT 780
 QY 2364 ATCACTTTTGTTCAAAAGTCGCTTATGAAATGGCGTGAAGAAAGAAAGCTGGGCTTAGAG 2423
 Db 781 GTTCATTTTGCAGGAGGCTGCTTGAAGAGGGGGTAAAGAACTATATGACTCAAG 840
 QY 2424 CGCCAAGTTTACCGGTCAAAAACCTGCGTAAACATCACCAAGAAAGACTTCAAGTTCAAC 2483
 Db 841 AAGAGGTTGTAGCGGTGTCCAACTGAGGCGAGCTCACAAAACCTCCACATGAAGTGAAT 900
 QY 2484 GACAAAACGGCAAAAATCACGCTGATCCGAAAACCTTCGAGTCTTTGTAGATGGCAA 2543
 Db 901 GATCGGCTTCCAGAGATCACCGTGGATCCAGAGACATACGTTGTTCACAGCAAAATGGCGAG 960
 QY 2544 CT-CTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCACGCGCTACACTTTCTTCT 2601
 Db 961 GTCTTAGCATGTGGCGCAGCCGATTAGTCCCTCTCTGCGCGCGGAACATTTCTCTCT 1019

RESULT 6
 DR637093
 LOCUS EST1027718 FvM Gibberella moniliformis cDNA clone FvMAT95, mRNA
 DEFINITION sequence.
 ACCESSION DR637093
 VERSION DR637093.1 GI:70711927
 KEYWORDS EST.
 SOURCE Gibberella moniliformis
 ORGANISM Gibberella moniliformis
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
 REFERENCE 1 (bases 1 to 946)
 AUTHORS Brown, D.W., Cheung, F., Proctor, R.H., Burchko, A.E., Zheng, L., Lee, Y.,
 Unterback, J., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D.,
 Kendra, D.F., Town, C.D. and Whitelaw, C.A.
 TITLE Comparative analysis of 87,000 expressed sequence tags from the
 fumonisin-producing fungus *Fusarium verticillioides*

JOURNAL Fungal Genet. Biol. 42 (10), 848-861 (2005)

PUBMED
COMMENT

Contact: Brown, D.W.
USDA/ARS/NCAUR

1815 N. University St, Peoria, IL 61604, USA

USA

Tel: 309 681 6230

Fax: 309 681 6689

Email: brown@ncaur.usda.gov

TIGR sequence name: FVMT95TH

Seq primer: AAT TAA CCC TCA AAG GG.

Location/Qualifiers

1..946

/organism="Gibberella moniliformis"

/mol_type="mRNA"

/strain="m3125"

/db_xref="taxon:117187"

/clone="FVMT95"

/tissue_type="mycelia"

/clone_lib="FVM"

/note="Vector: pBlueScript II SK(+); Site 1: EcoRI;

Site 2: XhoI; anamorph: Fusarium verticillioides. Library

FVM was prepared from pooled RNA obtained from a 48-hour

and a 72 hour, liquid GYM culture from strain M-3125.

Cultures were vacuum filtered and the mycelial mats were

frozen in liquid nitrogen, ground to a powder, and then

added to Trizol Reagent (Invitrogen, Carlsbad, CA) at

approximately 1 g mycelia per 10 ml Trizol. The cDNA was

directionally ligated into the pBlueScript II SK(+) XR

vector (cDNA Synthesis Kit; Stratagene)."

Db 542 ATCATCTCCGTGATAGCATCAAAATGTTCTGCCATCATGACCAACCTCAAGACCA 601
Qy 1809 TATACCAATTAATACGTTGCGAACAACATTTAGACATGCTCATGACATGCCACCATCTAGAC 1868
Db 602 TTCACACCAATACTCTCGATGAGCATCTCGATATGCTTAATGCTGCGCATCACTTGTCC 661
Qy 1869 AAACGCATCCGCGAGGATTTCAATTTCTCAAGCCGATATCCGCCCGGCTTATCGCG 1928
Db 662 AAGAATATCCAGAGGATGTAGCCTTCGCGAGAGCCGTTATTCGTCTGAAACCAATTGT 721
Qy 1929 GCTGAAGATGCTCCATGATATGGTGTGATGCGGATGACAAAGCTCGGATTCGCAAGCA 1988
Db 722 GCTGAGGATGTATTACACGACAAAGCGCTATCAGCATGATGAGCTCTGACTCGCAGGCT 781
Qy 1989 ATGGGGCTGCGAGCGCAAGTGAATCTCTGAACTTGGCAGACTGCGGATAGGAATAAAAAA 2048
Db 782 ATGGGCCGTTCGCGAGAGGTCTGTTTAAGAACATGATATGCGCATTAATAAATAGTG 841
Qy 2049 GAATTTGTAAGCTTCCTGAAGAT--GGCAAAGATAACGATAATTTCCGCATTAAGCGC 2105
Db 842 CAGAGGGCTTGGTTGCGGAGGATGAGGCGACAGGGGCTGATAATGCGCG-TGTAACGC 900
Qy 2106 TACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGCGCT 2149
Db 901 TATGTCAGCAAGTATATCTATACCACGAGCTATTTGCTCAGGGCTT 944

RESULT 7

DV133265

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Euphorbia esula (leafy spurge)

Euphorbia esula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

Rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;

Euphorbieae; Euphorbia.

1 (bases 1 to 809)

REFERENCE

AUTHORS

Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A.,

Kim, W.R. and Mikel, M.

Direct submission, Anderson, J.V. 2005

Unpublished (2005)

CONTACT: James V. Anderson

USDA/ARS

1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND

58105, USA

Tel: 701 239 1263

Fax: 701 239 1252

Email: andersjv@fargo.ars.usda.gov

Adaptors and tags in 5'-end sequenced clones:

(Vector) . . . TAACTTCATATCG (End Vector) (Start

EcoRI adaptor) AATTCATTCGTGTGG (End EcoRI adaptor) (Start

Insert) . . . AAAAAAAAAAAAAAAAAA (End Insert) (Start Tag) TGCCT (End

Tag) (Start NotI site/Vector) GCGCGCCACCCGCGG. . . Base

Calling/Quality Scores: PHRED from Washington University Genome

Center. Vector Trimming: Cross match from Washington University

Genome Center PHRAP suite. Low quality bases (Phred score < 20)

were trimmed from both ends of the sequence by an in-house script.

This sequence is vector free and at least 200 bp in length.

PCR Primers

FORWARD: TAATACGACTCACTATAGG (T7)

BACKWARD: ATTAACCCCTCACTAAAG (T3)

Insert Length: 809 Std Error: 0.00

Plate: CV03067A1 row: A column: 09

Seq primer: TAATACGACTCACTATAGG (T7)

High quality sequence stop: 809.

Location/Qualifiers

1..809

FEATURES

source

ORIGIN

Query Match 11.5%; Score 330; DB 10; Length 946;
Best Local Similarity 60.4%; Pred. No. 4.9e-67;
Matches 570; Conservative 0; Mismatches 370; Indels 4; Gaps 2;
Qy 1209 GATGGGTAAAGCCCTCATATGCTGGTGTGGGACAGAGCACTAGCAGGGGAAGGT 1268
Db 2 GACGCGGTAAAGAGGCGATGCTGCGAAGCTGACGAGTGTCTGACAGGTGAAGGA 61
Qy 1269 ATGATTATACCGTGGGGATGATTCACACACCACTTCTTCTCCACAACTTC 1328
Db 62 AAGATTGTGACCGGGGGCTATGACACGATATTCATTTATCTGCGCTCAGCAAGTA 121
Qy 1329 CCTACCGCTTAGCAATGGGCTTACACCACTGTTTGGAGCGGCACAGGTCCTGTAGAT 1388
Db 122 CCGAGCTCTGATCTGGTGTACCACTGTTGGCGTGTGACCGGCCAAGTCT 181
Qy 1389 GGCACGAATCGACTACTATCACTCCGGGCAAAATGGAATTTGCACCGCATGTGCGCGCA 1448
Db 182 GGAACGAACCACTACTTGTACGCTGCTCAATTACATGCGTCAAAATGTTGACGGCG 241
Qy 1449 GCAGAGATATTATGAATGTGGCTTTTGGGCAAGCAATAGCTTAGCAAAAAA 1508
Db 242 TSGCATCAGCTTCTTCAATTTGGTATTACTGCAAGGTAATGATGTTCTCTGAG 301
Qy 1509 CAATTTGAGAACAGTAGAGCGGCGGATTTGTTTAAATTCATGAAGACTGGGCGC 1568
Db 302 GGTCTGCCGATCAGTCAATGCTGTGCTGTGGCCCTCAAGCTTCATGAGCAGTGGGT 361
Qy 1569 ACAACACCAAGTGGGATCGATCACTGCTTGGCGTGGCAGATGAATACGATGTGCAAGTT 1628
Db 362 TGCATCTCTGCTGCTATTGACGCTTGTCTCAGTGTCTGTGATGAATTCGATTTCAATGT 421
Qy 1629 TGTATCCACCGGATCAGTCAATGAGGCGAGGTATGTAGATGACACCCCTAAATGCAATG 1688
Db 422 CTTATTACACTGACACCGCTTAAAGAGTCTGGCTTTTGTGCAATCTACGATCGCTGCTTC 481
Qy 1689 AACGGGCGCGCATCCATGCTTACACATTTGAGGCGGCTGGGAGGACACTCACTGAT 1748
Db 482 AAGAACCGCAAAATTTACTATTATACACAGAGGGTGCAGGAGGTGGCCATGCTCCGGAT 541
Qy 1749 GTTATCACCATGGGAGCGAGTCAATATTCTACCTCTCCACCCCACTATTCC 1808

Db	721	GC	AAGCATCCCAACACCCGACCGGTGATNTGAGCCCTATGTTTGGAGCATTTGGCAAG	720
Qy	2346	GC	AAATTTGACACACCGATCATCTTTTGT	2374
Db	781	GCT	GGAAGTGCTAACTCCATCTTTGT	809
RESULT 8				
DR633422				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
PUBLISHED				
COMMENT				
FEATURES				
source				
ORIGIN				
Query Match				
Best Local Similarity				
Matches				
Qy	1566	GC	CAACCAAGTCGATCGATCTGCTTGGAGGTGGCAGATCAATACGATGGCAA	1625
Db	1	GGA	ACTACTCTCTGTGTAGACAATCTGTTGAACGTAGCAGAAGATATGATCTGCAG	60
Qy	1626	GT	TGTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTAGATCACACCCCTAAATGCA	1685
Db	61	GT	TAAATATCACACGACACCTTGAAAGATCCGGAATTTGTGGACACAAATGCTGCA	120
Qy	1686	AT	GAACGGGCGCCATCCATGCTTACCAATTTAGGGAGCGGTGGAGACACTCACCT	1745
Db	121	TT	TAAAGGAGAACTATACATCTTATCAAGTGAAGTGCAGGAGTGTGATGCTCG	180
Qy	1746	GAT	GTATCACCATGCGAGGAGCTCAATATCTTACCTCTCTCCACACCCCGACTATT	1805
Db	181	GAC	ATTATAAAGTATGTGGTGTAAAAATGTTATACCGTCTACACGAATCTTACTCG	240
Qy	1806	CC	TATACCAATTAACGGTTGCAGAACCTTAGACATCTCATGACATGCCACACCTA	1865
Db	241	CC	TTTACTTCCAAATCAATGATGAACATCTCGATATCTGATGTTGCCATCATCTG	300
Qy	1866	GAC	AAACCGCATCCGAGGATTTACAAATTTCTCAAGCCGTATCCGCCCGGCTTATC	1925
Db	301	GAT	AAGACATTTCCGGAAGATGTAGCTTTTCGTGAGTCAGGTAAGGCCGGAACAATC	360
Qy	1926	GG	CGCTGAAGATGTCTCATGATATGGGTGTATCGCATGACAAAGCTCGGATTCGAA	1985
Db	361	GCT	GACAGAGACATATTCATGATATGGGGCTATTAGTATCATTTCTCTCTGATTTCCCA	420
Qy	1986	GCA	ATGGGGGTCAGGCGAAGTATCTCCGAATTCGACACTTCGGATCGGATAGAAATAA	2045
Db	421	GCT	ATGGGTCGATTTGGAGAGGTATATGACGAACCTTGGCAACTGCTCACAAAGATGA	480
Qy	2046	AA	AGAAATTTGGTAAGCTTCTCGAAGATGCGAAAGATAACGATAATTTCCGATTAAGCG	2105
Db	481	TC	ACAAAGGGATCAATTTGGTCCGGGCAATCAGATAAGATTAATTTCCGGATTAACGA	540
Qy	2106	TAC	ATCTCCAAATACACTATCAACCCCGCTTTTGACCCAGCGGTGAGCGAGTATATCGC	2165
Db	541	TAC	ATTCTTAAGTACACTATAAATCTCGATTAAGCAATGGGTTTGGCGATTAATTTGG	600
Qy	2166	TC	TGTGGAAGGCGCAAGATCCCGACTTGGTGGTGTGGAATCTTCCTTTTGGCGTA	2225
Db	601	TCA	GTGAGGTAGGAAATTTTCTGTATCTTGAATGTGGAAGCCCGCTTTCTTTGGGCA	660
Qy	2226	AAA	CCCAAAATCGTATCAAGCGCGTATGGTGGTCTTCTCTGAAATGGCGCATCTAAC	2285
Db	661	AA	CCGAAATGTTAAAGGAGGTGTAATTTGCTTGGCGAGACATGGGTGATCAAT	720
Qy	2286	GG	CTGTGTCCTCCCAACCGGTTTATATACCGCGAAATGTTTGGCGATCACGCGAAG	2345

```

QY 1492 ATAGCTCTAGCAAAAACAACTGTGTAGAACAGTAGAAGCGGCGCGATGTGTTTAAAT 1551
Db 190 ATGATAGATTCTCTGAGGGGTCTGCGCGATCAGGTCAATGCTGTGTGCTTGGCCCTCAAGC 249
QY 1552 TGCAATGAAGACTGGGGCAGACACCAAGTGGATCGATCACTGCTGAGCGTGCAGATG 1611
Db 250 TTCAATGAGGACTGGGGTGGCACTCTGCTGCTATTGAGCTTGTCTAGTGTCTGTGATG 309
QY 1612 AATACGATGTGCAAGTTTGTATCCACCCGATACAGTCAATGAGGCGAGTTATGATG 1671
Db 310 AATTCGATATTCAATGCTTATTACACTGACACCGCTTAACGAGTCTGGCTTGTCTGAAT 369
QY 1672 ACACCCTAAATGCAATGAAGCGGCGCGATCCATGCTTACCAATGAGGAGCGGGTG 1731
Db 370 CTACGATCGTGTCTTCAAGAACCGCAATTCATCTATATCAACAGAGGTGCGAGGAG 429
QY 1732 GAGGACACTCACCTGATGTTATACCACTGGCAGCGAGTCAATATCTACCTCCCTCCA 1791
Db 430 GTGGCCATGCTCCGATATCACTCTCGTGTGAGGATCAAAATGTTCTGCCATCATGA 489
QY 1792 CCACCCCTCACTATTCCCTATACCAATTAATACGGTTGAGAACACTTAGACATGCTCATGA 1851
Db 490 CCAACCTACAAGACCACTTACACGCAATCTCTCGATGACATCTCGATATGCTATGG 549
QY 1852 CATGCCACACTAGACAAACGATCCGCGAGGATTTACAATTTCTCAAGCGGTATCC 1911
Db 550 TCTGCCATCACTTGTCCAAGAAATATCCAGAGGATGTAGCCCTTCGCGAGAGCGGTATTC 609
QY 1912 GCCCGGCTCTATCGCGCTGAAGATGTCTCATGATATGCTGATGCTGATCGGATGACAA 1971
Db 610 GTGCTGAACCAITTGCTGCTGAGGATGATTACACGAAAGGCGCTATCAGCATGATGA 669
QY 1972 GCTCGGATTCGCAAGCAATGGGGGTGCGAGCGAAGTGAATTCCTCGAATTTGGCAGACTG 2031
Db 670 GCTCTGACTCGCAGCTATGGCGGCTTCCGAGAGGTGCTTTAAGAACATGGGATCTG 729
QY 2032 CGGATAGAATAAAAAAGAAATTTGGTAGCTTCTGAGAT---GGCAAGATTAACGATA 2088
Db 730 CGCAATAAAATAAGGTGCGAGAGGTTGGTTCGCGAGGATGAGGCGACAGGGGCTGATA 789
QY 2089 ATTTCCGATTAAGCGCTACATCTCCAAATACACTATCAACCCGCTTTGACCCACGGCG 2148
Db 790 ATGCGCGTGTAAACGCTATGCTAGCAAGTATATCTATTAACCCAGCTATTGCTCAGGGCT 849
QY 2149 TGACGAGTATATCGCTCTGTGAAGAGGCAAGATCGCGCA-CTTGGTGTGTGGAAT 2207
Db 850 TTGACATGTCATTTGGAGCAATGAGGTGGAAAGTTTGCTGATCTGTCTTTTGGAT 909
QY 2208 CTTGCCCTTTTGGC 2222
Db 910 CTTGCCGTGTTGGC 924

RESULT 9
COL122599
LOCUS GR_Eb04F01.f GR_Eb Gossypium raimondii cDNA clone GR_Eb04F01 5',
DEFINITION mRNA sequence.
ACCESSION COL122599.1 GI:48821286
VERSION EST.
KEYWORDS COL122599.1
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 872)
REFERENCE Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
AUTHORS Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing

```

```

FEATURES
    source             location/Qualifiers
    1..872
    /organism="Gossypium raimondii"
    /mol_type="mRNA"
    /db_xref="taxon:29730"
    /clone="GR_Eb04F01"
    /tissue_type="floral"
    /dev_stage="3 to +3 DPA"
    /lab_host="DH10B"
    /clone_lib="GR_Eb"
    /note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
    EcoRV; library made by invitrogen with RNA supplied by
    Wendle lab. Directional cloned into NotI-EV. Colonies
    plated/picked by AGI. More glycerol clones held in -80."

ORIGIN
    Query Match      10.7%; Score 309; DB 8; Length 872;
    Best Local Similarity 59.7%; Pred. No. 4.8e-62;
    Matches 519; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

QY 1260 GGGGAAGGTATGATTATTACCGCTGGGGGAATCGATTACACACCCACTTCTCTTCTCCA 1319
Db 2 GGAGAGAGTATGATCGTAACCTGACGGGCTATTGACTGCTGATGCTATTCATATGCTCT 61
QY 1320 CAACAATTCCTACCGCTTAGCCAAATGGCGTTACAACCATGTTTGGAGCGGCACAGGT 1379
Db 62 CAGTTGGTTTCATGAAGCTATTTCAGATGGCATCAACAATTTAGTTGGAGGTGGACCGGA 121
QY 1380 CTTGTAGATGCAAGCAATCGGACTACTATCCTCCGGGCAATGGAATGTCACCCGATG 1439
Db 122 CCAGCTGAAGGAACACGTCGAACCTACTTGTACGCCAGCTCCGTCGCAATGAAATGATG 181
QY 1440 TTGGCGCAGCAGAGAGTATTCTATGATGTGGGCTTTTGGGCAAAAGGCAATAGCTCT 1499
Db 182 CTCGAGTCCACTGATGACTTCCCTCTAAATTTTGGCTTCACAGGAGGAAATGTTCT 241
QY 1500 AGCAAAAACAACTTTGTAGAACAGTAGAAGCGGCGCGGATTTGGTTTAAATTTGCATGA 1559
Db 242 AAACCTGAAGAAATACATGAATAATCAAAGCGGAGCAATGGACTGAAACTGCAATGAG 301
QY 1560 GACTGGGCAACAACCAAGTCGATCTCATCTGCTTGAGCGTGGCAGATGAATAGAT 1619
Db 302 GATTGGGGAACCTACACCTCTGCAATAGACAGTTGTTGGCTGTTCGAGAACTATATGAT 361
QY 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCAAGTTATGTAGTACACCCCTA 1679
Db 362 ATCCAAGTTAATATTATCTACTGACACCTTTGAACGAATCTGGAATTTTGGAAACAAAT 421
QY 1680 AATGCAATGAAGCGGCGCGCCATCCATGCTACTACCACTTGAAGGAGCGGTGGAGACAC 1739
Db 422 AATGCAATTTAAGGAAGAACTATTACACTTATCAAGTGAAGGTGCGCGTGTGTCAT 481
QY 1740 TCACCTGATTTATCAACCAATGGCAGCGAGCTCAATATTCTACCTCTCCACCCCCC 1799
Db 482 GCTCAGATATCATCAAAGATGTGGCGTTAAAAACGTCCTCCCTTCGTCAACAAACCCG 541
QY 1800 ACTATTCCTTATACCATTAATACGGTTGCAGAACCTTAGACATGCTCATGATGATGCCAC 1859
Db 542 ACTGCCCTTTACTTCCAAATACTATAGATGAACATCTTGAATGCTGATGTTGTTGCCAT 601
QY 1860 CACCTAGCAAAACGCAATCCGCGAGGATTTACAATTTCTCAAGCGGTATCCGCCCGGC 1919
Db 602 CACCTCAGCAGGATATTCTCGAAGATGTAGCATTTGCAGAAATCAAGGATTAGGCGGGA 661
QY 1920 TCTATCGCGGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATGACAGCTCGGAT 1979

```

Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595
Fax: 520 621 1259

Email: rwing@genome.arizona.edu
Plate: 04 row: F column: 01.

location/Qualifiers
1..872

/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_Eb04F01"

/tissue_type="floral"

/dev_stage="3 to +3 DPA"

/lab_host="DH10B"

/clone_lib="GR_Eb"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:

EcoRV; library made by invitrogen with RNA supplied by

Wendle lab. Directional cloned into NotI-EV. Colonies

plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 10.7%; Score 309; DB 8; Length 872;

Best Local Similarity 59.7%; Pred. No. 4.8e-62;

Matches 519; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

QY 1260 GGGGAAGGTATGATTATTACCGCTGGGGGAATCGATTACACACCCACTTCTCTTCTCCA 1319

Db 2 GGAGAGAGTATGATCGTAACCTGACGGGCTATTGACTGCTGATGCTATTCATATGCTCT 61

QY 1320 CAACAATTCCTACCGCTTAGCCAAATGGCGTTACAACCATGTTTGGAGCGGCACAGGT 1379

Db 62 CAGTTGGTTTCATGAAGCTATTTCAGATGGCATCAACAATTTAGTTGGAGGTGGACCGGA 121

QY 1380 CTTGTAGATGCAAGCAATCGGACTACTATCCTCCGGGCAATGGAATGTCACCCGATG 1439

Db 122 CCAGCTGAAGGAACACGTCGAACCTACTTGTACGCCAGCTCCGTCGCAATGAAATGATG 181

QY 1440 TTGGCGCAGCAGAGAGTATTCTATGATGTGGGCTTTTGGGCAAAAGGCAATAGCTCT 1499

Db 182 CTCGAGTCCACTGATGACTTCCCTCTAAATTTTGGCTTCACAGGAGGAAATGTTCT 241

QY 1500 AGCAAAAACAACTTTGTAGAACAGTAGAAGCGGCGCGGATTTGGTTTAAATTTGCATGA 1559

Db 242 AAACCTGAAGAAATACATGAATAATCAAAGCGGAGCAATGGACTGAAACTGCAATGAG 301

QY 1560 GACTGGGCAACAACCAAGTCGATCTCATCTGCTTGAGCGTGGCAGATGAATAGAT 1619

Db 302 GATTGGGGAACCTACACCTCTGCAATAGACAGTTGTTGGCTGTTCGAGAACTATATGAT 361

QY 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCAAGTTATGTAGTACACCCCTA 1679

Db 362 ATCCAAGTTAATATTATCTACTGACACCTTTGAACGAATCTGGAATTTTGGAAACAAAT 421

QY 1680 AATGCAATGAAGCGGCGCGCCATCCATGCTACTACCACTTGAAGGAGCGGTGGAGACAC 1739

Db 422 AATGCAATTTAAGGAAGAACTATTACACTTATCAAGTGAAGGTGCGCGTGTGTCAT 481

QY 1740 TCACCTGATTTATCAACCAATGGCAGCGAGCTCAATATTCTACCTCTCCACCCCCC 1799

Db 482 GCTCAGATATCATCAAAGATGTGGCGTTAAAAACGTCCTCCCTTCGTCAACAAACCCG 541

QY 1800 ACTATTCCTTATACCATTAATACGGTTGCAGAACCTTAGACATGCTCATGATGATGCCAC 1859

Db 542 ACTGCCCTTTACTTCCAAATACTATAGATGAACATCTTGAATGCTGATGTTGTTGCCAT 601

QY 1860 CACCTAGCAAAACGCAATCCGCGAGGATTTACAATTTCTCAAGCGGTATCCGCCCGGC 1919

Db 602 CACCTCAGCAGGATATTCTCGAAGATGTAGCATTTGCAGAAATCAAGGATTAGGCGGGA 661

QY 1920 TCTATCGCGGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATGACAGCTCGGAT 1979

```

Db      662  ACCATTGCTGCCGAAGACATATTGATGACACAGGGCAATCAGCATTTATCTTCAGAT 721
Qy      1980  TCGAAGCAATGCGGGGTGCGAGGCAAGTATTCCTCGAATTCGCGACAGCTGCGGATTAAG 2039
Db      722  TCAAGCGCATGCGGTGCGATTCGCGAGGTGATATGACAACTTGGCAAAACGCCCCACAAG 781
Qy      2040  AATAAAAAGAAATTTGGTAAGCTTCCTGAAATGCGCAAGATAGCAAGTAATTTCCGCATT 2099
Db      782  ATGAATATCAAAAGGGGACCACTTGGTCTCGAGACTCAGACAAACGCAATCTAAGAAATC 841
Qy      2100  AAGCGCTACATCTCCAAATACACTATCA 2128
Db      842  AAAGTTATATAGCAAAATACACCATAAA 870

RESULT 10
DV759139
LOCUS   DV759139      889 bp      mRNA      linear      EST 22-NOV-2005
DEFINITION PchrSEQ10171 Phanerochaete chrysosporium pBluescript (EcoRI-XhoI)
Phanerochaete chrysosporium cDNA clone PchrSEQ10171, mRNA sequence.
ACCESSION DV759139
VERSION   DV759139.1 GI:82603228
KEYWORDS EST.
SOURCE    Phanerochaete chrysosporium (anamorph: Sporotrichum pruinosum)
ORGANISM Phanerochaete chrysosporium
          Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
          Aphyllophorales; Corticiaceae; Phanerochaete.
REFERENCE 1 (bases 1 to 889)
AUTHORS   Tsang, A., Storms, R. and Butler, G.
TITLE     Expressed sequence tags from Phanerochaete chrysosporium cDNA
          library
JOURNAL   Unpublished (2005)
COMMENT   Contact: Tsang A
          Centre for Structural and Functional Genomics
          Concordia University
          7141 Sherbrooke Street West, SP545-1, Montreal, QC H4B1R6, Canada
          Tel: 514 848 2424 3405
          Fax: 514 848 4504
          Email: tsang@vax2.concordia.ca
          POLYA=No. Location/Qualifiers
FEATURES             source
     source           1..889
                     /organism="Phanerochaete chrysosporium"
                     /mol_type="mRNA"
                     /db_xref="taxon:5306"
                     /clone="PchrSEQ10171"
                     /lab_host="E. coli"
                     /clone_lib="Phanerochaete chrysosporium pBluescript
                     (EcoRI-XhoI)"
                     /note="Vector: pBluescript KS+; Site 1: XhoI; Site 2:
                     EcoRI; ded, double-stranded cDNA. The EcoRI-XhoI-digested
                     cDNA was ligated with EcoRI-XhoI-digested pBluescript KS+
                     (Invitrogen Corp). "
ORIGIN
Query Match      10.7%; Score 306.8; DB 10; Length 889;
Best Local Similarity 60.3%; Pred. No. 1.6e-61;
Matches 536; Conservative 0; Mismatches 347; Indels 6; Gaps 1;
Qy      1338  CTAGCCCAATGCGGTTACAAACCATGTTTGGAGCGGCGCACAGTCTCTGTAGATGGCAACGAAT 1397
Db      1   CTCGCGCGGGCGACGACGACGATGATCGCGCGGCGACGCGCGCGTCCAGCGGCACGAC 60
Qy      1398  GCGACTACTATCATCTCGGGCAATAGAACTTGACCGCATGTTGCGGCGAGCAGAGAG 1457
Db      61  GCGACGACGTGCACGCGCGACCGCGTCTTACATGCGCCACATGCTCGCTGCGACAGCGG 120
Qy      1458  TATTCTATGAATGGGCTTTTGGGCAAGGCAATAGCTCTACCAAAAACACTTTGTA 1517
Db      121  CTGCGGATGAACCTTGGCTTTACGGGCAAGGCAACGACTCGGGGTACGGCGCCTGGAG 180
Qy      1518  GAACAAGTAGAAGCGGCGCGATTTGGTTTAAATTTGATGAAGACTGGGGCACAACACCA 1577

```

```

Db      181  GAGATCGTGCAGCGCGGCATCGGTCTCAAGCTCCAGAGGACTGGGGCAACAACGCCG 240
Qy      1578  ACTGCGATCGATCACTGCTTGGAGCGTGGCAGATGAATAGATGCGAAGTTTGTATCCAC 1637
Db      241  TCAGCCATCATCAGCTGTCTTGACGTGCGAGACAAGTAGCAGCGTGCAGGTCAATATTCAT 300
Qy      1638  ACCGATACAGTCAATGAGCGCAGTTATGTAGATGACACCTTAATCAATGAACGCGGCGC 1697
Db      301  ACGGATACGCTCAACGAGAGTGGATTTGTGTAGAGCACTATCAAGGCATTTCGGAAATCGC 360
Qy      1698  GCATCCATGCTTACCAATTTCTCAAAGCCGTATCCGCCCGCGCTCTATCGCGGCTGAAGAT 1757
Db      361  ACCATTCAACCTATCAACACGAGGGTGCAGTGGCGGACACGCACTGATATCATCATCACC 420
Qy      1758  ATGCGAGGCGAGCTCAATTTCTCAAGCCCTCTCCACCAACCCCACTATTCCTTATACCAT 1817
Db      421  GTCTCGGAATTCACAACAGTACTGCTCATCGTCCACGAACCCCAACGCGCCATACGCAAAAC 480
Qy      1818  AATACGGTTGCAAGACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCATC 1877
Db      481  AACACACTGGACGAGCATCTCGACATGCTCATGGTGTGCCACCACTCGACAAGAACATC 540
Qy      1878  CGCGAGGATTTACAATTTTCTCAAAGCCGTATCCGCCCGCGCTCTATCGCGGCTGAAGAT 1937
Db      541  CGGAGGACATCGCGTTGCGAGATGCGCGCATCCGCCGCGAGACCGTCCGCCGCGAGGAC 600
Qy      1938  GTGCTCCATGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGCT 1997
Db      601  GTGCTGACGACATCGCGCGCATCTCGATGATCTCGTCCGACTTCGCGAGCGATGGGCGCG 660
Qy      1998  GAGCGCAAGTATTCCTCGAACTTTGGCGAGACTGCGGATAAGATAAAAGAAATTTGGT 2057
Db      661  GTCGCGGAGGTGCTGCGCGCGACGTGGCGCACCGCGCAAGATGCGCGAGTTTGGCGGC 720
Qy      2058  AAGCT-----TCCTGAAGATGCGCAAGATAAGATAATTTCCGCATTAAAGCCCTACATC 2111
Db      721  CCGCTCGCGCGGCTCGCGGACGTCGAGGCTCGCAACACGCGCGCTGAAGCGGTATCATC 780
Qy      2112  TCCAAATACACTATCAACCCCGCTTTGACCCACGCGCTGAGCGAGTATATCGGCTCTGTG 2171
Db      781  TCCAAGTACAGATCAACCGCGGATCAGCATGGATGAGCCACCTCGTTGACAGETC 840
Qy      2172  GAAGAGGCAAGATCGCGACTTGTGTGTGGAATTCCTGCCTTTTTT 2220
Db      841  GCGGTGCGGGCGGCTCGCGACCTCGTCTCTGGAAGCCGCGAGAACTTTG 889

RESULT 11
CX676965
LOCUS   CX676965      748 bp      mRNA      linear      EST 19-JAN-2005
DEFINITION ydd429h04.y1 Sea urchin EST Lib1 Strongylocentrotus purpuratus cDNA
clone ydd29h04 5' similar to TR:Q9SR52 Q9SR52 PUTATIVE UREASE. ;,
mRNA sequence.
ACCESSION CX676965
VERSION   CX676965.1 GI:57936920
KEYWORDS EST.
SOURCE    Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
          Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
          Echinoidea; Euechinoidea; Echinacea; Echinoida;
          Strongylocentrotidae; Strongylocentrotus.
REFERENCE 1 (bases 1 to 748)
AUTHORS   Coffman, J.A., Robertson, A.J., Clifton, S., Pape, D., Hillier, L.,
          Martin, J., Wylie, T., Dante, M., Meyer, R., Theising, B., Bowers, Y.,
          Gibbons, M., Ronko, I., Tsagareishvili, R., Ritter, E., Kennedy, S. and
          Wilson, R.
          WashU Sea Urchin EST Project
          Unpublished (2004)
          Contact: Dr. James A. Coffman
          WashU Sea Urchin EST Project
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
          Tel: 314 286 1800

```

Fax: 314 286 1810
 Email: est@watson.wustl.edu
 DNA sequencing by: Washington University Genome Sequencing Center
 Seq primer: -28RppOT
 High quality sequence stop: 680.
 Location/Qualifiers
 source
 1..748
 /organism="Strongylocentrotus purpuratus"
 /mol_type="mRNA"
 /db_xref="taxon:7668"
 /clone="ydd29h04"
 /lab_host="DH10B"
 /clone_lib="Sea urchin EST Lib1"
 /note="Vector: pCMVSPORT6.1 (Invitrogen); Site 1: NotI;
 Site 2: SmaI; Arrayed normalized library of full-length
 cDNAs representing blastula stage transcriptome of the sea
 urchin Strongylocentrotus purpuratus, cloned into the
 vector pCMVSPORT6.1 (Invitrogen)"

FEATURES
 source
 1..748
 /organism="Strongylocentrotus purpuratus"
 /mol_type="mRNA"
 /db_xref="taxon:7668"
 /clone="ydd29h04"
 /lab_host="DH10B"
 /clone_lib="Sea urchin EST Lib1"
 /note="Vector: pCMVSPORT6.1 (Invitrogen); Site 1: NotI;
 Site 2: SmaI; Arrayed normalized library of full-length
 cDNAs representing blastula stage transcriptome of the sea
 urchin Strongylocentrotus purpuratus, cloned into the
 vector pCMVSPORT6.1 (Invitrogen)"

ORIGIN
 Query Match 10.5%; Score 302.2; DB 9; Length 748;
 Best Local Similarity 63.1%; Pred. No. 1.9e-60;
 Matches 466; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
 1648 TCAATGAGGAGGTATGTAGATGACACCCCTAAATGCAATGAAGCGGCGGCCATCCATG 1707
 9 TCAATGAATCTGCTGTGGAAGATACATAGTGTCTTTGAGGGCGCCACCATCCACA 68
 1708 CCTACCACTTGGAGGCGGTGGAGGACACTCACCTGATGTATACCATGCGAGCG 1767
 69 CCTATCACTCGAGGGTGCAGAGAGGCCACGCCCTTGACATCATGAAGTAGTGGCG 128
 1768 AGCTCAATATTCTACCTCTCCACCCGCCACTATTCCCTATACCAATTAATAGGTTG 1827
 129 TGCCCAATGTATACCTCTGCTACCAACCAACCGGCCCTTCACATGACACCATCG 188
 1828 CAGAACCTTAGACATGTCTATGATGACGACACCTAGACAAAGCATCCGAGGATT 1887
 189 ATGAACACCTGGATATGCTATGTGCTGCCACCATCTTGACAAGAAATCTAAGGAAGACG 248
 1888 TACAATTTCTCAAGCGGTATCCGCCGCTCTATCGCGCTGAAGATGTCTCCATG 1947
 249 TTGCTTTGACAGGTCTCGCATCAGACCGCAACCAATGCTGCTGAAGATATCTTCAATG 308
 1948 ATATGGGTGTGATCGCATGATCAAGCTCGATTTCGAAGCAATGGGGGTGCGAGCGAAG 2007
 309 ATCTGGGAGCAATCAGTATTGGCTCCGATTTCGAGGCCATGGGGGTGTCGGTGAGG 368
 2008 TGATTCTCGAATTTGGCAGACTCGGATAGAAATATAAAGAAATTTGGTAACTTCTCTG 2067
 369 TGATCACCGGTACATGGCAGACGGCAGATAAGATGAAGATTTTCAGAGGCGAGATGTG 428
 2068 AAGATGGAAGATAACGATATTTCCGATTAAGCGGTACATCTCCAAATACACTATCA 2127
 429 AGSAAACGGGTGACAAACGACACCACTTAAGGGTCAAGCGTTACATTGCCAAGTACAGGTCA 488
 2128 ACCCGCTTTTCAACCCAGCGGTGACGAGTATATCGGCTCTGTGGAAGAGGCGAAGATCG 2187
 489 ACCGGCAATGGCCACGGCATGGGCCATCTCATTTGGCTCTGTACAGGTTGGTAAATGG 548
 2188 CCGACTTGGTGTGGAAATCTGCTTTTGGGTAAACCCCAAAATCGTGATCAAG 2247
 549 CTGACCTGTTTCTTGGAAATCGGCATTTCTCGGAGCTAAACCTGACCTCATCAAG 608
 2248 GCGGTATGGTGTCTTCTGAAATGGGGGATTTAAACGGTCTGTGGCCACTCCCAAC 2307
 609 GAGGTACATATGCTTTGGGCACAAATGGGCGATGCCCAATGCATCCATCCGACCCCTGAAC 668
 2308 CGGTTTATTACCGCAATATTTGGGCATCAGGCAAGCGGAAATTTTCACACCCAGCATCA 2367
 669 CTGTCAAGACAGGAAATATTTGGTGTCTTACGGGAAGTCCATCGTGAGAACTCTGTCA 728
 2368 CTTTGTGTTTCCAAAGTCGC 2386

Db 729 TCTTTGTCTCAAGCAGC 747
 RESULT 12
 LOCUS DR637386
 DEFINITION EST1028011 FvM Gibberella moniliformis cDNA clone FVMAV95, mRNA
 ACCESSION DR637386
 VERSION DR637386.1 GI:70712220
 KEYWORDS EST
 SOURCE Gibberella moniliformis
 ORGANISM Gibberella moniliformis
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
 REFERENCE 1 (bases 1 to 899)
 AUTHORS Brown,D.W., Cheung,F., Proctor,R.H., Butchko,A.E., Zheng,L., Lee,Y.,
 Utterback,T., Smith,S., Feidblyum,I., Glenn,A.E., Plattner,R.D.,
 Kendra,D.F., Town,C.D. and Whitelaw,C.A.
 Comparative analysis of 87,000 expressed sequence tags from the
 fungus-producing fungus Fusarium verticillioides
 Fungal Genet. Biol. 42 (10), 848-861 (2005)
 16099185
 CONTACT: Brown, D.W.
 USDA/ARS/NCAR
 1815 N. University St, Peoria, IL 61604, USA
 Tel: 309 681 6230
 Fax: 309 681 6689
 Email: brown@ncar.usda.gov
 TIGR sequence name: FVMAV95TH
 Seq primer: AAT TAA CCC TCA CTA AAG GG.
 FEATURES
 Location/Qualifiers
 1..899
 /organism="Gibberella moniliformis"
 /mol_type="mRNA"
 /strain="m3125"
 /db_xref="taxon:117187"
 /clone="FVMAV95"
 /tissue_type="mycelia"
 /clone_lib="FvM"
 /note="Vector: pBlueScript II SK(+) XR; Site 1: EcoRI;
 Site 2: XhoI; anamorph: Fusarium verticillioides. Library
 FvM was prepared from pooled RNA obtained from a 48-hour
 and a 72 hour, liquid GVAM culture from strain M-3125.
 Cultures were vacuum filtered and the mycelial mats were
 frozen in liquid nitrogen, ground to a powder, and then
 added to Trizol Reagent (Invitrogen, Carlsbad CA) at
 approximately 1 g mycelia per 10 ml Trizol. The cDNA was
 directionally ligated into the pBlueScript II SK(+) XR
 vector (CDNA Synthesis Kit; Stratagene)."
 ORIGIN
 Query Match 10.4%; Score 300.2; DB 10; Length 889;
 Best Local Similarity 59.5%; Pred. No. 5.9e-60;
 Matches 524; Conservative 0; Mismatches 353; Indels 3; Gaps 1;
 1312 TTTCTCCACAAATTCCTACCGCTCTAGCCAAATGGCGTTACACCATGTTGGAGCG 1371
 10 TCTGCCCTCAGCAAGTACCCGAAGCTCTTGCACTCTGGTAAACCACTGCTTGGCGGT 69
 1372 GCACAGGTCTCTAGATGGCAGAAATGCGACTACTATCACTCCGGCAAAATGGAATTCG 1431
 70 GTACCGGCCAAGTCTGGAACGAAACGCACTACTGTAGCGCTGTGCTCATTCATGC 129
 1432 ACCGATGTTGGCGCAGCAGAAAGATTTCTATGAATGGGCTTTTGGGCAAGGCA 1491
 130 GTCAAAATGTTGAGCGGTGCGATCAGCTTCCTATCAATATGGTATTTACTGGCAAGGTA 189
 1492 ATAGCTTAGCAAAAACAACTTGTAGAAACAAGTAGAAGCGGCGCATGTTTAAAT 1551
 190 ATGATAGTTCTCTGAGGGTCTGCGCGATCAGGTCAATGCTGTGCTTGTGCTCTCAAGC 249

	QY	1552	TGATGAAGACTGGGGCAACAACCAAGTCCGATCGAATCACTGTAGCGTGCGCAGATG	1611
	Dd	250	TTCATGAGGACTGGGGTTGCACCTCCTGCTATTGACGCTTGCTCAGTGTCTGTGATG	309
	QY	1612	AATACGATGCGAAGTTTGTATCCACACCGATACAGTCANTGAGGCCAGGTTATGTAGATG	1671
	Dd	310	AA'TTGGATATTCAATGTCTTA'TT'CACTACACGCTTAA'CGAGTCTGGCTT'TGT'CGAAT	369
	QY	1672	ACACCCTAAATGCAATGAACGGCGGGCCCATCGCTTACCACATTTGAGGGAGCGGGTG	1731
	Dd	370	CTAGGATCGCTGCTTTCAGAACCGCACATTCATACTTATCACACAGGGTGCAGGAG	429
	QY	1732	GAGGACACTCACCTGATGTTATCCATGCGAGCGAGCTCAATATTCTACCTCTCCCA	1791
	Dd	430	GTGGCCATGCTCCGGATATCATCTCCGTGTGAGCATCAAATGTTCTGCCATCATCGA	489
	QY	1792	CGACCCCCTACTTCCCTATACCA'TTAATACGGTTGCAGNACHACTTAGACATGCTCATGA	1851
	Dd	490	CCAACCTTCAAGACCAITTCACACGCAATACTCTCGATGAGCATCTCGATATGCTTATGG	549
	QY	1852	CATGCCACACCTAGACAAAACGATCCGGGAGGATTTACAA'TTTCT'CAAAGCCGTATCC	1911
	Dd	550	TCTGCCATCACTTGTTCCAAGAATATCCGAGAGATGTAGCCTTCGGCGGAGACCGGTATTC	609
	QY	1912	GCCCCGGCTCTATCGCGGCTGAAGATGTGCTCCATGATATGGTGATCGCGATGACAA	1971
	Dd	610	GTGCTGA AACCATTCGTCTGAGGATGTATTACGACAAAGSCGTATCAGCATGATGA	669
	QY	1972	GCTCGGATTTGCMAGCAATGGGGCGTGCGAGCGAAGTGATTCCTCGAACTTGGCAGACTG	2031
	Dd	670	GCTCTGACTCGAGGCTATGGGCCGTTGGCGGAGGTCGTTTTAAGAACATGGAATACTG	729
	QY	2032	CGSATAAGATAAAAAAGAA'TTGGTAAGCTTCCTCAAGT---GCCAAGATAACGATA	2088
	Dd	730	CGCATAAAATTAAGTGCAGAGGGTTGGTTGCGGAGGATGAGGGCACAGGGGCTGATA	789
	QY	2089	ATTTCGCGCTTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTTGACCCACGGCG	2148
	Dd	790	ATGCGGCTGTA A A A C G T A T G T C A G C A A G T A T A T T A A C C A G C T A T T G C T C A G G G C T	849
	QY	2149	TGAGCGAGTATATCGGCTCTGTGGAAAGAGCGCAAGATCGC	2188
	Dd	850	TTGGCATGTCATTCGAAGCATTTGAGTTTGGAAAGTTTGC	889

RESULT 13

DR633939

LOCUS

DEFINITION

11

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISMI

REFERENCE

AUTHORS

11

37111

TO: DIRECTOR

JOURNAL
DIMMED

COMMENT
PUBLISHED

COMMENT

```

TIGR sequence name: FWMA846TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.
Location/Qualifiers
1. .796
  /organism="Gibberella moniliformis"
  /mol_type="mRNA"
  /strain="m3125"
  /db_xref="taxon.117187"
  /clone="FWMA846"
  /tissue_type="mycelia"
  /clone_lib="FvM"
  /notes="Vector: pBlueScript II SK(+) XR; Site 1: EcoRI;
  Site 2: XhoI; anamorph: Fusarium verticillioides. Library
  FvM was prepared from pooled RNA obtained from a 48-hour
  and a 72 hour, liquid GYAM culture from strain M-3125.
  Cultures were vacuum filtered and the mycelial mats were
  frozen in liquid nitrogen, ground to a powder, and then
  added to Trizol Reagent (Invitrogen, Carlsbad CA) at
  approximately 1 g mycelia per 10 ml Trizol. The cDNA was
  directionally ligated into the pBlueScript II SK(+) XR
  vector (CDNA Synthesis Kit; Stratagene)."

```

vecc01 (CDNA Synthesis Kit, Sciencelab.com).									
ORIGIN	Query Match	10.3%;	Score 296.4;	DB 10;	Length 796;				
	Best Local Similarity	60.8%;	Pred. No. 4.6e-59;						
	Matches 493; Conservative	0;	Mismatches 311;	Indels	0;	Gaps	0;		
QY	1243	GCACAGACCTAGCAGGGGAAGGTATGATTAATTACCGCTGGGGGAATCGATTCAACA	1302						
DB	3	GCACGGATGCTGATGAGGTGAAGGAAGATTGTGACCGCGGCGCTATCGACACGCATA	62						
QY	1303	CCCACCTCTTCTCCACAACAATTCCTTACCCTCTAGCCCAATGGCGTTACACCACTGT	1362						
DB	63	TTCCACTTATCTGCCCTCAGCAAGTACCCGAAGCTCTTGCACTGTGGTGTAAACCACTGC	122						
QY	1363	TTGGAGGCGCACAGGTCCTGTAGATGGCACGAATGCGACTACTATCACTCCGGGCAAT	1422						
DB	123	TTGGCGGTGTACCGGCCCAAGTCTGGAAAGAAACGCACTACTTGTACGCCTGGTCTC	182						
QY	1423	GGAACTTGCACCGCATGTTGCGCGCAGCAGAAGATATCTATGAATGTGGGCTTTTGG	1482						
DB	183	ATTACATCGCTCAAAATGTTTGAGCGGTGCGATCAGCTTCTATCAATATTGGTATTACTG	242						
QY	1483	GCAAGGCATATAGCTTAGCAAAAACAATTGTGTAGAACAAGTAGAAGCGGCGCGATTG	1542						
DB	243	GCAAGGTAATATAGTTCTCTGAGGGTCTGCGCGATCAGGTCAATGCTGGTCTTGTG	302						
QY	1543	GTTTTAAATGTCATGAAGACTGGGGCACACCAAGTGCATCGATCACTGTCTTGACG	1602						
DB	303	GCCTCAAGCTTCATGAGSACTGGGGTTGCACTCCTGCTATTTGACGCTTGTCTCAGTG	362						
QY	1603	TGSCAGATCAATACGATGTGCAAGTTTGTATCCACACCGCATACAGTCAATGAGGCAGTT	1662						
DB	363	TCGTGATGAATTCGATATCAATGTCTTATTCACACTGCACGCTTAACGAGTCTGGCT	422						
QY	1663	ATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCCATCCATGCTACCACATTCGAGG	1722						
DB	423	TTGTGAACTACGATCGCTGCTTTCAAGAACCGCACAAATTCATCTTATCACACAGAG	482						
QY	1723	GAGCGGCTGGAGACACTCACCTGATGTTATACCAATGCGACGGGAGCTCAATATTCTAC	1782						
DB	483	GTGACGAGGTGCCATGCTCCGGATATCATCTCCGTGGTAGAGCATCAAAATGTTCTGCG	542						
QY	1783	CTCTCTCCACACCCCACTATTCCTATACCAATTAATACGGTTGCAGACACTTAGACA	1842						
DB	543	CATCATCGAACCAACCTACAAGCAATTCACAGCAATACTCTCGATGAGCATCTCGATA	602						
QY	1843	TGCTATGACATGCCACCACTAGACAAACGGATCCGCGAGGATTTACAATTTTCTCAA	1902						
DB	603	TGCTTATGCTTGCCATCACTTGTCCAGAAATATCCAGAGGATGTAGCCCTTCGCGAGA	662						
QY	1903	GCCGTATCCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCATGATATGGGTGTGATCG	1962						

```

Db      663  GCGTATTGCTGCTGAACCAATTGCTGCTGAGGATGTTATTACAGCAAAAGCGCGTATCA 722
Qy      1963  CGATGACAAGCTCGGATTGCGAAGCAATGGGCGGTGCGAGCGGAAGTGAATCTCGAACTT 2022
Db      723  GCATGATGAGCTCTGACTCGCAGGCTATGGCGGTGCGGAGAGGTGCTTTTAAGAACAAT 782
Qy      2023  GGCAGACTGCGGAT 2036
Db      783  GGAATACTGCGCAT 796

RESULT 14
DR635599
LOCUS      EST1026224 FvM Gibberella moniliformis cDNA clone FVMAJ96, mRNA
DEFINITION
ACCESSION  DR635599
VERSION    DR635599.1 GI:70710433
KEYWORDS   Gibberella moniliformis
SOURCE     Gibberella moniliformis
ORGANISM   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE  1 (bases 1 to 849)
AUTHORS    Brown,D.W., Cheung,F., Proctor,R.H., Butchko,A.E., Zheng,L, Lee,Y.,
            Uterback,T., Smith,S., Feldblyum,T., Glenn,A.E., Plattner,R.D.,
            Kendra,D.F., Town,C.D. and Whitelaw,C.A.
TITLE      Comparative analysis of 87,000 expressed sequence tags from the
            funomisin-producing fungus Fusarium verticillioides
JOURNAL    Fungal Genet. Biol. 42 (10), 848-861 (2005)
PUBMED     16099185
COMMENT    Contact: Brown, D.W.
            USDA/ARS/NCAUR
            1815 N. University St, Peoria, IL 61604, USA
            Tel: 309 681 6230
            Fax: 309 681 6689
            Email: brown@ncaur.usda.gov
            TIGR sequence name: FVMAJ96TH
Seq primer: AAT TAA CCC TCA AAG GG.
FEATURES
            source
            1..849
            /organism="Gibberella moniliformis"
            /mol_type="mRNA"
            /strain="m3125"
            /db_xref="taxon:117187"
            /clone="FVMAJ96"
            /tissue_type="mycelia"
            /clone_lib="FvM"
            /note="Vector: pBlueScript II SK(+); Site 1: EcoRI;
            Site 2: XhoI; anamorph: Fusarium verticillioides. Library
            FvM was prepared from pooled RNA obtained from a 48-hour
            and a 72 hour, liquid GYM culture from strain M-3125.
            Cultures were vacuum filtered and the mycelial mats were
            frozen in liquid nitrogen, ground to a powder, and then
            added to Trizol Reagent (Invitrogen, Carlsbad CA) at
            approximately 1 g mycelia per 10 ml Trizol. The cDNA was
            directionally ligated into the pBlueScript II SK(+) XR
            vector (cDNA Synthesis Kit; Stratagene)."
```

Query Match 10.2%; Score 293; DB 10; Length 849;

Best local similarity 60.2%; Pred No. 3e-58;

Matches 503; Conservative 0; Mismatches 330; Indels 3; Gaps 1;

Qy 1317 CCACAACAAATCCCTACCGCTCTAGCCAAATGGCGTTACAAACATGTTGGAGGCGGCACA 1376

Db 3 CCTCAGCAAGTACCGGAAGCTCTTCATCTGCTCTAAACCCACCATGCTTGGCGGTGTACC 62

Qy 1377 GGTCTGTAGATGGCAAGATCGGACTACTATCATCTCCGGGCAATGGAACHTTGCACCGC 1436

Db 63 GGCCCAAGTGTGGAAACGAACCAACTACTTTGTACGCTGTGCTCATTTACATGCGTCAA 122

```

Qy      1437  ATGTTGGCGGAGCAGAGAGTATTCTATGAATGTGGGCTTTTGGGCAAGCAATAGC 1496
Db      123  ATGTTGCGAGGCTGGCGATCAGCTTCTCTATCAATATTGGTATTATCTGCGCAAGGTAATGAT 182
Qy      1497  TCTAGCAAAAACAACCTTGTAGAACAAAGTAGAAGCGGCGCGATTGGTTTAAATTGCAT 1556
Db      183  AGTTCTCTGAGGGTCTGGCGGATCAGGTCAATGCTGGTGTCTGAGTCTCTGTGATGAATTC 242
Qy      1557  GAAAGACTGGGCGACAAACCAAGTGCATCGATCACTGCTTGAGCGGTGGCAGATGAATAC 1616
Db      243  GAGGACTGGGTTGCACCTCTGCTGCTATTGAGCGTTGCTCAGTCTCTGTGATGAATTC 302
Qy      1617  GATGTGGAAGTTTGTATCCACCGATACAGTCAATGAGGCGAGTTATGTAGATGACACC 1676
Db      303  GATATTCAATGTCTTATTACACTGACACGCTTAACGAGTCTGGCTTTGTGCAATCTACG 362
Qy      1677  CTAAATGCAATGAACGGCGGCCCATCCATGCTACCACTGAGGAGCGGTGGAGGA 1736
Db      363  ATCGCTGCTTTCAAGAACCGCAAAATTCATCTTATCACAGAGGTTGAGAGGTGGC 422
Qy      1737  CACTCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTCTACCTCTCCACCAACC 1796
Db      423  CATGCTCCGGATATCATCTCCGTGGTAGAGCATCAAAATGTTCTGCCATCATCGACCAAC 482
Qy      1797  CCCATTTCCCTATACCAATTAATAGGTTGCAGAACACTTAGACATGCTCATGACATGC 1856
Db      483  CCTACAAGACCAATTCACAGCAATATCTCTCGATGAGCATCTCGATATGCTTATGGTCTGC 542
Qy      1857  CACCACTTAGCAAAACGGATCCGCGAGGATTTACAAATTTTCTCAAGCCGTATCCGCCCC 1916
Db      543  CATCACTTGTCCAAGAAATATCCAGAGGATGTAGCCCTCCGGAGAGCGGTATTCGTGCT 602
Qy      1917  GGCTCTATCGCGGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATGACAAAGCTCG 1976
Db      603  GAAACCATGCTGCTGAGGATGTATTACAGCAAGAGCGCTATCAGCATGATGAGCTCT 662
Qy      1977  GATTCGCAAGCAATGGGCGTGCAGCGAAGTATTCCTCGAACTTTGGCAGACTGCGGAT 2036
Db      663  GACTCGCAGGCTATGGCGGTTGCGGAGAGGTGCTTTTAAAGAACATGGAATATCTCGCAT 722
Qy      2037  AGAATAAAAAGAAATTTGGTAAGCTTCTCTGAGAT---GGCAAGATAACGATAATTTC 2093
Db      723  AAAATAAAGGTGCAGAGGGTTGGTTGCCGAGGATGAGGGGCACAGGGGCTGATAATGCG 782
Qy      2094  CGCATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTTGACCCACCGCGT 2149
Db      783  CGGTAAACGCTATGCTAGCAAGTATATTAATTAACCCAGCTATTGCTCAGGGCTT 838
```

RESULT 15

AZ935182/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AZ935182 784 bp DNA linear GSS 24-APR-2001
 BJ_Ba003020f B. japonicum BAC library Bradyrhizobium japonicum
 genomic, genomic survey sequence.

AZ935182
 AZ935182.1 GI:13777494
 GSS.

Bradyrhizobium japonicum
 Bradyrhizobium japonicum
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 Bradyrhizobiaceae; Bradyrhizobium.

1 (bases 1 to 784)
 Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A.,
 Goicoechea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.

A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
 genome
 Genome Res. 11 (8), 1434-1440 (2001)

11483585
 Contact: Wing RA
 Clemson University Genomics Institute

Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288

Fax: 864 656 4293
 Email: rwing@clmson.edu
 Class: BAC ends
 High quality sequence stop: 719.
 Location/Qualifiers
 1..784
 /organism="Bradyrhizobium japonicum"
 /mol_type="genomic DNA"
 /strain="USD110"
 /db_xref="taxon:375"
 /lab_host="E. coli"
 /clone_lib="B. japonicum BAC library"
 /note="Vector: pindigo536; Site_1: HindIII"

FEATURES
 source

ORIGIN

Query Match	9.9%;	Score 284;	DB 11;	Length 784;
Best Local Similarity	64.0%;	Pred. No. 4.1e-56;		
Matches 458;	Conservative 0;	Mismatches 253;	Indels 5;	Gaps 3;

Qy	1468	ATGTGGGCTTTTGGGCAAGCAATAGCTCTAGCAAAAACAACTTTAGAACAAAGTAG	1527
Db	715	ATCTCGGCAATTCGGGCAAGGCAACGCTCGCGCCCGCGCTGGTTCGAGATCATCA	656
Qy	1528	AAGCGGGCGGATGGTTTAAATTGCATGAGACTGGGGCACACCAAGTGGATCG	1587
Db	655	AGGCGGGCGCATGGCGCTGAAGCTGCACGAGGATTTGGGGCAC-ACGCGCGCGCGATCG	597
Qy	1588	ATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAG	1647
Db	596	ACAATGCTGTCTGGTGGCCGACGATACGATCCAGGTCTATGATCCACCGATACGC	537
Qy	1648	TCAATGAGGCAAGTTATAGATGACACCCCTAAATGCAATGAACGGGGCGGCCATCCATG	1707
Db	536	TGACGAATCGGGCTTCGTCGAGGATACGATCAAGGCGTTCAAGGGCGCACCATCCACG	477
Qy	1708	CCTACCAATTCAGGAGCGGGTGGAGACACTCACTGATGTTATACCATGGCAGGG	1767
Db	476	CTTTCCACACCGAGGGCGCGCGCGGTACGCCCGCGACATCAAGGTTCGAGGGC	417
Qy	1768	AGCTCAATATTTACCTCTCCACCCACCTATTTCCCTATATACCATTAATACGGTTG	1827
Db	416	TGAAGACGTGCTGCGGTCATCGAACACCGCGCGCCCTTACCCGCAACCATCG	357
Qy	1828	CAGAACATTCAGACATGCTCATGATCATGCCACCTAGACAAACGCAATCCCGAGGATT	1887
Db	356	ACGACATTCGACATGCTGATGTTGTCACCACTCGATCCCTCGATCGCGAAGATC	297
Qy	1888	TACAAATTTCTCAAGCGGTATCCGCCCGCGCTTATCGCGGCTGAAGATGTGCTCCATG	1947
Db	296	TTGGGTTCCGCGAAGCGGTATCCGCAAGGAGACCATCGCGCGCGAGGACATCCTGCACG	237
Qy	1948	ATATGGGTGTCATCGCGATGACAGCTCGGATTCGCAAGCAATGGG-CGCTCAGGGCGAA	2006
Db	236	ATCTCGGCGCGCTTCGATGATGTCCTCGNACTCCAGGCCATGGGCGCGCTCGGGCGAA	177
Qy	2007	GTGATTCCTCGAACTTGGCAGACTCGGATGAAGATAAAGAAATTTGGTAAGCTTCC-	2065
Db	176	GTCAATCCCGACTTGGCAGACCGCGCAAGATGAAGAGCAGCGCGGATCGCTGCGG	117
Qy	2066	--TGAAGATGCCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACT	2123
Db	116	CAGGACAAAGGCAAGGACAAAGCAATTTCCGCGTCAAGCGCTACATCGCCAAATACAG	57
Qy	2124	ATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGG	2179
Db	56	ATCAACNCCGATCGCGCACGGGTTGCGAAGCTGATCGGTTTCGGTGGAGAGGG	1

Search completed: August 10, 2006, 05:05:16
 Job time : 12173 secs

Db	1	VKLTPEQEKFLYYAGEVARKKAEGLKLNQPPAIAYISAHIMDEARRGKKTVAELMEE	60
Qy	61	CMHFLKKDEVMGPGVNNVPDLGVEATFPDGTGLVTNNWPIEPDHFHFKAGEVKFGCDKDIE	120
Db	61	CMHFLKKDEVMGPGVNNVPDLGVEATFPDGTGLVTNNWPIEPDHFHFKAGEVKFGCDKDIE	120
Qy	121	LNAGKEVTELEVNTNEGPKSLHVGSHFIPFFPAANKALKFDREKAYGKRLDIPSGNTLIRIGAG	180
Db	121	LNAGKEVTELEVNTNEGPKSLHVGSHFIPFFPAANKALKFDREKAYGKRLDIPSGNTLIRIGAG	180
Qy	181	QTRKVQLIPLGGSKKVIIGMGLVNNIADERRHKKALDKAKSHGFIK	226
Db	181	QTRKVQLIPLGGSKKVIIGMGLVNNIADERRHKKALDKAKSHGFIK	226
RESULT 3			
ADJ58241			
ID	ADJ58241	standard; protein; 226 AA.	
XX	ADJ58241;		
XX	06-MAY-2004	(first entry)	
DT	XX		
DE	XX	UreaseXY subunit #1.	
KW	XX	immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.	
OS	XX	Helicobacter felis.	
PN	XX	EP1176192-A2.	
XX	XX		
PD	XX	30-JAN-2002.	
PF	XX		
PF	XX	11-JUL-2001; 2001EP-00202666.	
XX	XX		
PR	XX	17-JUL-2000; 2000EP-00202565.	
XX	XX		
PA	XX	(ALKU) AKZO NOBEL NV.	
XX	XX		
PI	XX	Kusters JG, Cattoli G;	
XX	XX		
DR	XX	WPI; 2002-124384/17.	
DR	XX	N-PSDB; ADJ58240.	
XX	XX		
PT	PT	Novel Helicobacter felis urease X and Y subunit polypeptides, useful in	
PT	PT	the diagnosis of Helicobacter felis infections and in the preparation of	
PT	PT	vaccines.	
XX	XX		
PS	PS	Disclosure; SEQ ID NO 5; 76pp; English.	
XX	XX		
CC	CC	The present invention relates to a novel Helicobacter felis urease X and	
CC	CC	Y subunit polypeptides and immunogenic fragments. The polypeptides are	
CC	CC	used in the manufacture of vaccines against Helicobacter felis infections	
CC	CC	and in diagnostic tests to detect antibodies against Helicobacter felis.	
CC	CC	Helicobacter felis is difficult to grow so it is more convenient to use	
CC	CC	the expression products of the genes encoding the urease X and Y subunits	
CC	CC	in the manufacture of vaccines. The present sequence represents an	
CC	CC	ureaseXY subunit of the invention.	
XX	XX		
SQ	XX	Sequence 226 AA;	
Query Match			
Best Local Similarity 99.4%; Score 1174; DB 5; Length 226;			
Matches 224; Conservative 1; Mismatches 1; Indels 0; Gaps 0			
Qy	1	VKLTPEQEKFLYYAGEVARKKAEGLKLNQPPAIAYISAHIMDEARRGKKTVAELMEE	60
Db	1	VKLTPEQEKFLYYAGEVARKKAEGLKLNQPPAIAYISAHIMDEARRGKKTVAELMEE	60
Qy	61	CMHFLKKDEVMGPGVNNVPDLGVEATFPDGTGLVTNNWPIEPDHFHFKAGEVKFGCDKDIE	120
Db	61	CMHFLKKDEVMGPGVNNVPDLGVEATFPDGTGLVTNNWPIEPDHFHFKAGEVKFGCDKDIE	120

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
QY 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226

RESULT 4
ADJ58247
ID ADJ58247 standard; protein; 226 AA.
XX
AC ADJ58247;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit #5.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EPI1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
DR N-PSDB; ADJ58246.
XX
Novel Helicobacter felis urease X and Y subunit polypeptides, useful in the diagnosis of Helicobacter felis infections and in the preparation of vaccines.
XX
PS Disclosure; SEQ ID NO 11; 76pp; English.
XX
The present invention relates to a novel Helicobacter felis urease X and Y subunit polypeptides and immunogenic fragments. The polypeptides are used in the manufacture of vaccines against Helicobacter felis infections and in diagnostic tests to detect antibodies against Helicobacter felis. Helicobacter felis is difficult to grow so it is more convenient to use the expression products of the genes encoding the urease X and Y subunits in the manufacture of vaccines. The present sequence represents an ureaseXY subunit of the invention.

QY Sequence 226 AA;
Query Match 99.4%; Score 1174; DB 5; Length 226;
Best Local Similarity 99.1%; Pred. No. 4.5e-45;
Matches 224; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGVAVARRKKAEGKLQNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGVAVARRKKAEGKLQNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDHFKAGEVKGCDKDIE 120
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDHFKAGEVKGCDKDIE 120
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
QY 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226

Db 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226

RESULT 5
ADJ58244
ID ADJ58244 standard; protein; 226 AA.
XX
AC ADJ58244;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit #3.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EPI1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
DR N-PSDB; ADJ58243.
XX
Novel Helicobacter felis urease X and Y subunit polypeptides, useful in the diagnosis of Helicobacter felis infections and in the preparation of vaccines.
XX
PS Disclosure; SEQ ID NO 8; 76pp; English.
XX
The present invention relates to a novel Helicobacter felis urease X and Y subunit polypeptides and immunogenic fragments. The polypeptides are used in the manufacture of vaccines against Helicobacter felis infections and in diagnostic tests to detect antibodies against Helicobacter felis. Helicobacter felis is difficult to grow so it is more convenient to use the expression products of the genes encoding the urease X and Y subunits in the manufacture of vaccines. The present sequence represents an ureaseXY subunit of the invention.

QY Sequence 226 AA;
Query Match 99.1%; Score 1170; DB 5; Length 226;
Best Local Similarity 98.7%; Pred. No. 6.8e-45;
Matches 223; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGVAVARRKKAEGKLQNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGVAVARRKKAEGKLQNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDHFKAGEVKGCDKDIE 120
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDHFKAGEVKGCDKDIE 120
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
QY 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226

RESULT 6
ADQ37849
ID ADQ37849 standard; protein; 234 AA.

XX AC ADQ37849;
XX DT 07-OCT-2004 (first entry)
XX DE H. bizzozeronii urea polypeptide.
XX KW Urease; urease gene cluster; urease structural gene;
XX KW urease accessory gene; ureA; Helicobacter bizzozeronii infection;
XX KW antibacterial; enzyme.
XX OS Helicobacter bizzozeronii.
XX PN US2004142343-A1.
XX PD 22-JUL-2004.
XX PF 12-AUG-2003; 2003US-00639273.
XX PR 16-AUG-2002; 2002US-0404337P.
XX PA (CHAN/) CHANG Y.
XX PA (SIMP/) SIMPSON K W.
XX PA (ZHUJ/) ZHU J.
XX PI Chang Y, Simpson KW, Zhu J;
XX WPI; 2004-533502/51.
XX DR N-PSDB; ADQ37848.
XX PT Novel isolated nucleic acid molecule having urease gene cluster, and
XX FT conferring on Helicobacter bizzozeronii ability to produce urease, useful
XX PT as vaccine for preventing disease in mammals infected by H.bizzozeronii.
XX PS Claim 23; SEQ ID NO 3; 40pp; English.
XX CC The invention relates to an isolated nucleic acid molecule conferring on
XX CC Helicobacter bizzozeronii an ability to produce urease, where the nucleic
XX CC acid molecule is a urease gene cluster comprising at least one urease
XX CC structural gene and at least one urease accessory gene. The nucleic acid
XX CC molecule is chosen from ureA, ureB, ureE, ureF, ureG, ureH and ureI. The
XX CC invention also relates to an isolated protein encoded by the nucleic
XX CC acid, a vaccine for preventing onset of disease in mammals infected by H.
XX CC bizzozeronii comprising a nucleic acid and a carrier, and an isolated
XX CC antibody or its binding portion raised against the nucleic acid. The
XX CC nucleic acids, proteins and antibodies are useful for vaccinating mammals
XX CC against onset of disease caused by infection of H. bizzozeronii, which
XX CC involves administering the sequences. The sequences are useful for
XX CC detecting H. bizzozeronii in a sample of tissue or body fluids which
XX CC involves providing a nucleic acid as an antigen, providing an antibody,
XX CC or providing a nucleotide sequence as a probe in a nucleic acid
XX CC hybridisation assay, contacting the sample with the antigen or the probe,
XX CC and detecting any reaction which indicates that H. bizzozeronii is
XX CC present in the sample. This sequence represents the H. bizzozeronii ureA
XX CC polypeptide.
XX SQ Sequence 234 AA;
Query Match 56.3%; Score 665; DB 8; Length 234;
Best Local Similarity 57.3%; Pred. No. 1.2e-21;
Matches 130; Conservative 39; Mismatches 52; Indels 6; Gaps 3;
Qy 1 VKLTPKEDEKFLLYAGVARKRAEGKLNQPEAIVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKLMLHYAGLAKRANGVKLNTTEVALISAHVMEERAGKKSVDLMOE 60
Qy 61 CMHFLKKDEVMFPGVGNMVPDLGVEATFPDGTGLVTNNPIEPDEHFKA-GEV--KFGCDK 117
Db 61 GRTLKADDDVMPGVAHMEHVEGIEANFPDGTGLVTIHTPVEDGGHKLAPGEVILK--NE 117
Qy 118 DIENAGKEVTELVNTPGKSLHVGSHFFPEANKALKFDEKAYGKRLDIPSGNTURI 177
Db 118 DITLNAGQATTTLEVHNKGRDPVQVQSGSHFFPEVNNKLEFDEKAYGKRLDIASGTAVRF 177

Qy 178 GAGQTRKVLQIPLGGSKKVIQMGVGNLNNIADRRHKKHAKLDKAKSHGF 224
Db 178 EPGEKKTVELIQIGGNRIYGFNSLVDRQADTDGKKLAKRAKERGF 224

RESULT 7

AAU35875
ID AAU35875 standard; protein; 238 AA.

XX AC AAU35875;

XX DT 14-FEB-2002 (first entry)

XX DE Helicobacter pylori cellular proliferation protein #188.

XX KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX KW antibacterial; drug design.

XX OS Helicobacter pylori.

XX PN WO200170955-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US009180.

XX PR 21-MAR-2000; 2000US-0191078P.

XX PR 23-MAY-2000; 2000US-0206848P.

XX PR 26-MAY-2000; 2000US-0207272P.

XX PR 23-OCT-2000; 2000US-0242578P.

XX PR 27-NOV-2000; 2000US-0253625P.

XX PR 22-DEC-2000; 2000US-0257931P.

XX PR 16-FEB-2001; 2001US-0269308P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;

XX DR WPI: 2001-611495/70.

XX DR N-PSDB; AAS53734.

XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.

XX PS Example 3; SEQ ID NO 11468; 51pp; English.

XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes
XX CC themselves and the encoded proteins. The prokaryotes used are Escherichia
XX CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
XX CC antibodies capable of binding to the expressed proteins. The proteins can
XX CC be used to screen compounds in rational drug discovery programmes. The
XX CC antisense nucleic acid sequence is also useful to screen for homologous
XX CC nucleic acids which are required for cell proliferation in a wide variety
XX CC of organisms. The present sequence represents an essential prokaryotic
XX CC cellular proliferation protein. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 238 AA;

Query Match 52.4%; Score 619; DB 4; Length 238;

Best Local Similarity 53.1%; Pred. No. 1.6e-19;

Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKELLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
 Db 1 MKLTPKELDKMLHYAGELAKRKKGKLNVEVALISAHIMEARAGKKTAAELMQE 60

QY 61 CMHFLKKDEVMGPNVGDILGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
 Db 61 GRTLKPDVMDGVASMIHEVGIEAMPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
 Db 120 INEGKKAHSVKNVGNDRPVQIGSHFHFPEVNRCLDFREKTFGKRLDIASGTAVRPEPG 179

QY 181 QTRKVQLIPLGSKKVGIMGNGLVNNIADERHKHAKLDKAKSHGF 224
 Db 180 EEKVELIDIGNRRIFGFNALVDROADNESKIALHRAKERGF 223

RESULT 8
 AAR04579
 ID AAR04579 standard; protein; 238 AA.
 XX
 AC AAR04579;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 21-SEP-1990 (first entry)
 XX
 DE Part of protein with urease activity.
 XX
 KW Urease; probe.
 XX
 OS Helicobacter pylori.
 XX
 PN WO9004030-A.
 XX
 PD 19-APR-1990.
 XX
 PF 06-OCT-1988; 88FR-00013135.
 XX
 PR 06-OCT-1988; 88FR-00013135.
 XX
 PA (INSP) INST PASTEUR.
 PA (INRM) INSERM INST NAT SANTE & RECH MED.
 XX
 PI Labigne A;
 XX
 DR WPI; 1990-147844/19.
 DR N-PSDB; AAQ04328.
 XX
 XX New nucleotide sequences encoding Campylobacter pylori-ureaseants - and
 PT derived vectors, transformants, protein, antibodies and probes, useful in
 PT diagnosis, treatment and prevention of infections.
 XX
 PS Claim 11; Page 34; 47pp; French.
 XX
 CC The protein can be used for the prodn. of antibodies and to prepare
 CC vaccines for the prevention /treatment of C. pylori infections. See also
 CC AAR04578-81. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27
 CC -AUG-2003 to correct OS field.)
 XX
 SQ Sequence 238 AA;

Query Match 52.2%; Score 616; DB 2; Length 238;
 Best Local Similarity 52.7%; Pred. No. 2.2e-19;
 Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKELLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
 Db 1 MKLTPKELDKMLHYAGELAKRKKGKLNVEVALISAHIMEARAGKKTAAELMQE 60

QY 61 CMHFLKKDEVMGPNVGDILGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
 Db 61 GRTLKPDVMDGVASMIHEVGIEAMPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
 Db 120 INEGKKAHSVKNVGNDRPVQIGSHFHFPEVNRCLDFREKTFGKRLDIASGTAVRPEPG 179

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
 Db 120 INEGKKAHSVKNVGNDRPVQIGSHFHFPEVNRCLDFREKTFGKRLDIASGTAVRPEPG 179

QY 181 QTRKVQLIPLGSKKVGIMGNGLVNNIADERHKHAKLDKAKSHGF 224
 Db 180 EEKVELIDIGNRRIFGFNALVDROADNESKIALHRAKERGF 223

RESULT 9
 AAW07193
 ID AAW07193 standard; protein; 238 AA.
 XX
 AC AAW07193;
 XX
 DT 16-OCT-2003 (revised)
 DT 11-FEB-1997 (first entry)
 XX
 DE H. pylori urease A subunit.
 XX
 KW Urease; ureA gene; ureB gene; vaccine.
 XX
 OS Helicobacter pylori; strain CPM630.
 XX
 PN WO9633732-A1.
 XX
 PD 31-OCT-1996.
 XX
 PF 25-APR-1996; 96WO-US005800.
 XX
 PR 28-APR-1995; 95US-00431041.
 PR 06-DEC-1995; 95US-00568122.
 XX
 PA (ORAV-) ORAVAX INC.
 XX
 PI Lee CK, Monath TP, Ackerman SK, Thomas WD, Soman G, Kleanthous H;
 PI Weltzin RA, Pappo J, Ermak T, Guirakhoo F, Bhagat H, Suesman I;
 XX
 DR WPI; 1996-497373/49.
 DR N-PSDB; AAT44351.
 XX
 PT Vaccine for inducing mucosal response to Helicobacter - contg. multimeric
 PT urease complex and pref. an antibiotic, anti-secretory agent or bismuth
 PT salt.
 XX
 PS Disclosure; Page 70-71; 98pp; English.
 XX
 CC Urease A (AAW07193) and B (AAW07194) subunits of Helicobacter pylori are
 CC encoded by the ureA + ureB gene locus of clinical isolate CPM630.
 CC Vectors, e.g. pORV214 (see also AAT44351), carrying the gene locus can be
 CC used to transform host cells for the large-scale prodn. of recombinant,
 CC enzymatically inactive, multimeric urease. The urease complex induces a
 CC mucosal immune response that can treat or prevent Helicobacter, esp. H.
 CC pylori, gastroduodenal infection. (Updated on 16-OCT-2003 to standardise
 CC OS field)
 XX
 SQ Sequence 238 AA;

Query Match 52.2%; Score 616; DB 2; Length 238;
 Best Local Similarity 52.7%; Pred. No. 2.2e-19;
 Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKELLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
 Db 1 MKLTPKELDKMLHYAGELAKRKKGKLNVEVALISAHIMEARAGKKTAAELMQE 60

QY 61 CMHFLKKDEVMGPNVGDILGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
 Db 61 GRTLKPDVMDGVASMIHEVGIEAMPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
 Db 120 INEGKKAHSVKNVGNDRPVQIGSHFHFPEVNRCLDFREKTFGKRLDIASGTAVRPEPG 179

Db 120 INEGKKAVSVKVNVDPRPVOIGSHFFHFEVNRNCLDFDEKTFGKRLLDIAGTAVRFPG 224

Qy 181 QTRKVQLPIGGSKVKTCMGNLVNNIADERHKHALDKAKSHGF 224
: : ||| :||::: | | | : | | : | | |
Db 180 EEKSVELIDIGENRIFGFNALVDROQDNESKKTALHPAERGPF 223

RESULT 10
AAU35693
ID ID AAU35693 standard; protein; 238 AA.
XX AC
XX AAU35693;
AC XX
XX 14-FEB-2002 (first entry)
XX Helicobacter pylori cellular proliferation protein #6.
DE DE
XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
KW XX
XX Helicobacter pylori.
OS OS
XX WO200170955-A2.
PN PN
XX 27-SEP-2001.
PD PD
XX 21-MAR-2001; 2001WO-US009180.
XX PF
XX 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX PA
XX (ELIT-) ELITRA PHARM INC.
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR N-PSDB; AAS3552.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 11286; 51lpp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes,
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 238 AA;

Qy	1	VKLTPKEQKFLYYAGEVARKRKAAGLKLNOPEAIYISAHIMDEARRGKKTVAQLMBE	60
Ds	1	MKLTPKELDKLMLHYAGELAKRKEGKIKLNVYEAVALISAHIMEEARAGKKTAAELMQE	60
Qy	61	CMHFLKDEVMGPGVGNMVPDLGVZATPDGDKLVTWNWPIEPDEHFKADEVKFGCKDIE	120
Ds	61	GRTLLKDDVDMDGVSAMITHEVGIEBAMFPDGGKLVTVHTPIEANGKLVPQEL-FLKNEDIT	119
Qy	121	LNACKEVTELEVTVEGPKSLHVGSHFHFEEANKALKPDREKAYGKRLLDIPSGNTLIRIGAG	180
Ds	120	INEGKAVSVKVMVGDRPVQIGSHFHFEEVNRCLDFDREKTFGKRLLDASGTAVRFBEG	179
Qy	181	QTRKVVQLIPLGSGSKYIGMGNGLVNNIADERHKHKALDKAKSHGF	224
Ds	180	EKSVELIDIGENRRITFGFNALVDRQADNESKKTALHEAKERGF	223
RESULT	11		
AA	49644		
AA	49644	standard; protein; 238 AA.	
AC	49644;		
XX			
DT			
XX			
XX		17-MAY-2002 (first entry)	
DE		Urease A protein PA4.	
XX			
KW		Urease A; vaccine; T cell-stimulating peptide; antibacterial; antiulcer;	
KW		MHC; major histocompatibility complex; epitope; antiinflammatory;	
KW		cytostatic; virucide; protozoicide; immunosuppressive; pathogen; tumour;	
KW		autoimmune disease; Helicobacter pylori infection.	
XX			
OS		Unidentified.	
XX			
FN		WO200202141-A2.	
XX			
PD		10-JAN-2002.	
XX			
PF		04-JUL-2001; 2001WO-EP007656.	
XX			
PR		05-JUL-2000; 2000DE-01032538.	
XX			
PA		(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.	
XX			
PI		Aebischer A, Bumann D, Lucas B, Meyer TF;	
XX			
DR		WPI; 2002-164497/21.	
XX			
PT		Preparing vaccine containing T cell-stimulating peptides, useful	
PT		particularly for protection against Helicobacter pylori, by selecting	
PT		epitopes from target proteins.	
XX			
PS		Example 1; Fig 3; 36pp; German.	
XX			
CC		This invention describes a novel method for the preparation of a vaccine	
CC		based on T cell-stimulating peptide sequences. The method comprises (1)	
CC		identifying potential MHC (major histocompatibility complex)-binding and	
CC		T cell-reactive epitopes in a target protein; (2) eliminating cross-	
CC		reactive epitopes, and (3) verifying remaining epitopes by determining	
CC		their activity in vitro and/or in vivo. The products of the invention	
CC		have antibacterial, antiulcer, antiinflammatory, cytostatic, virucide,	
CC		protozoicide and immunosuppressive activity. The vaccine is especially	
CC		used to protect against infection by Helicobacter pylori, but more	
CC		generally can be prepared from proteins from any pathogen (virus,	
CC		bacterium, protozoan) or tumour, or associated with autoimmune disease.	
CC		Synthetic epitopes that are more immunogenic than natural sequences can	
CC		be identified, and the vaccine may have increased immunogenicity (when	
CC		formulated with adjuvant) or immunotolerance (formulated without an	
CC		adjuvant). This sequence represents the urease A protein PA4 described in	
CC		the invention	
XX			
SQ		Sequence 238 AA;	

Query Match 52.2%; Score 616; DB 5; Length 238;
 Best Local Similarity 52.7%; Pred. No. 2.2e-19;
 Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEOKFLLYAGVAVARRKKAEGKLNQPEAIAVISAHIMDEARRGKKTVAOLMEE 60
 DB 1 MKLTPKELDKLMLHYAGELAKRKEKGKILNYVEAVALISAHIMEARAGKKTAAELMQE 60

QY 61 CMHFLKDEVMGPNVMDLGVATFPDGTGLVTVMNPIEPDHFHFKAGEVKGCDKDIE 120
 DB 61 GRTLKPDVMDGVASMIHEVGIEAMPDGTGLVTVHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
 DB 120 INEGKAVSVKVNVDGPRVQIGSHFFHFEVNRCLDFDREKTFGKRLDIASGTAVRFPFG 179

QY 181 QTRKVLQIPLGSKKVIQMGVLNVIADERHKKALDKAKSHGF 224
 DB 180 EEKSVELIDIGNRRIFGFNALVDROADNESKIALHRAKERGF 223

RESULT 12
 AEB55115
 ID AEB55115 standard; protein; 245 AA.
 XX
 AC AEB55115;
 DT 22-SEP-2005 (first entry)
 XX
 DE Hylicobacter pylori Urase A protein fragment SEQ ID 4.
 XX
 KW pHUR3; vaccine; helicobacter pylori infection; urease;
 KW immune stimulation; antibacterial; immunostimulant; antiulcer;
 KW cytosolic; antinflammatory; gastrointestinal-gen.; gastritis;
 KW peptic ulcer; carcinoma.
 XX
 OS Helicobacter pylori.
 OS Salmonella typhimurium.
 XX
 PN W0200132014-A2.
 XX
 PD 10-MAY-2001.
 XX
 XX 01-NOV-2000; 2000WO-US030191.
 XX
 XX 01-NOV-1999; 99US-00431705.
 PR
 XX (ORAV-) ORAVAX INC.
 XX
 XX Kleanthous H, Londono-Arcila P, Freeman D;
 PI
 XX WPI; 2001-343379/36.
 DR
 DR N-PSDB; AEB55112, AEB55130.
 XX
 XX Inducing an immune response against Helicobacter in mammals, useful for
 PT treating Helicobacter induced gastroduodenal diseases.
 XX
 XX Disclosure; SEQ ID NO 4; 63pp; English.

The invention relates to inducing an immune response against Helicobacter in a mammal, comprising mucosally administering to the mammal an attenuated Salmonella vector containing a nucleic acid molecule encoding a Helicobacter antigen, and parenterally administering to the mammal a Helicobacter antigen. Also included is an attenuated Salmonella vector comprising a nucleic acid molecule encoding a Helicobacter antigen. The Helicobacter antigen is a urease subunit, or its immunogenic fragment (encoded by the ureA and ureB genes). The mammal is at risk of having but does not have Helicobacter infection or has a Helicobacter infection. The attenuated Salmonella vector further comprises an htrA or nlrB promoter. The vector can be used in inducing an immune response against Helicobacter in a mammal. The vector can be used to treat Helicobacter infection. The vector and the method can be used to treat

CC Helicobacter induced gastroduodenal diseases, including acute, chronic or atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal ulcers), and gastric carcinoma. Attenuated Salmonella vector pHUR3 was constructed expressing UreA and UreB and including an E. coli AmpR gene. The present sequence represents a protein or peptide encoded by the pHUR3 plasmid. NOTE: pHUR3 is represented by both AEB55112 and AEB55130, the peptides expressed by pHUR3 are shown in the sequence listing to be split between AEB55112 and AEB55130, yet Figure 4 shows all the peptides being expressed by the one sequence, therefore all encoded peptides are cross-referenced to both sequences.

XX
 XX Sequence 245 AA;
 SQ

Query Match 52.2%; Score 616; DB 4; Length 245;
 Best Local Similarity 52.7%; Pred. No. 2.4e-19;
 Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEOKFLLYAGVAVARRKKAEGKLNQPEAIAVISAHIMDEARRGKKTVAOLMEE 60
 DB 8 MKLTPKELDKLMLHYAGELAKRKEKGKILNYVEAVALISAHIMEARAGKKTAAELMQE 67

QY 61 CMHFLKDEVMGPNVMDLGVATFPDGTGLVTVMNPIEPDHFHFKAGEVKGCDKDIE 120
 DB 68 GRTLKPDVMDGVASMIHEVGIEAMPDGTGLVTVHTPIEANGKLVPGEL-FLKNEDIT 126

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
 DB 127 INEGKAVSVKVNVDGPRVQIGSHFFHFEVNRCLDFDREKTFGKRLDIASGTAVRFPFG 186

QY 181 QTRKVLQIPLGSKKVIQMGVLNVIADERHKKALDKAKSHGF 224
 DB 187 EEKSVELIDIGNRRIFGFNALVDROADNESKIALHRAKERGF 230

RESULT 13
 ADS09178
 ID ADS09178 standard; protein; 238 AA.
 XX
 AC ADS09178;
 XX
 DT 16-DEC-2004 (first entry)
 XX
 DE H. pylori urease alpha beta subunit.
 XX
 KW Urease; alpha subunit; UreB; gastric ulcer; peptic ulcer; gastritis;
 KW stomach cancer; vaccine; antibody; immune reaction.
 XX
 OS Helicobacter pylori.
 XX
 PN JP2004261080-A.
 XX
 PD 24-SEP-2004.
 XX
 XX 28-FEB-2003; 2003JP-00054654.
 PP
 XX 28-FEB-2003; 2003JP-00054654.
 PR
 XX (DOKU-) DOKURITSU GYOSEI HOJIN KAGAKU GIJUTSU SH.
 PA (FUKU-) FUKUYAMA RINSHO KENSA CENT KK.
 XX
 XX WPI; 2004-665475/65.
 DR
 XX Protein having specific region of beta-subunit urease of Helicobacter
 PT pylori, which generates antibodies that suppress proliferation and urease
 PT activity of bacteria, useful for preventing Helicobacter pylori
 PT infection.
 XX
 PS Disclosure; SEQ ID NO 3; 26pp; Japanese.
 XX
 CC The invention relates to a protein having a specific region of primary
 CC structure of the beta-subunit Helicobacter pylori urease, appearing as
 CC ADS09176 (amino acids 201-338 of the beta subunit) or a sequence which
 CC generates antibodies that suppresses proliferation of H. pylori and which

CC field.)

XX SQ Sequence 238 AA;

Query Match 51.5%; Score 608; DB 2; Length 238;
Best Local Similarity 52.2%; Pred. No. 5.2e-19;
Matches 117; Conservative 43; Mismatches 63; Indels 1; Gaps 1;

Qy	1	VKLTPKQEKFLYYAGVARRKKAEGKLNQPEAIAIYISAHIMDEARRGKKTVAQLMEE	60
Db	1	MKLTPKELDKLMLHYAGSLAKRKEKGIKLNVEAVALISAHIMEARAGKKTAAELMQE	60
Qy	61	CMHFLKKDEVMPGCVGMVDPDLGVEATPDDGTYKLVTVNWPIDPDEHFKAAGEVKFGCDKDIE	120
Db	61	GRTLLKPDVNDMGVASMIEHVGIEAMFDDGTYKLVTVNWPIDPDEHFKAAGEVKFGCDKDIE	119
Qy	121	LNAGKEVTELEVTEGPKSLHVGSHFHFPEANKALKPDREKAYGKRLDIPSGNTLRLIGAG	180
Db	120	INEGKAVSVKKNVGDPRVQIGSHFHFPEVNRCLDPREKTFHKKRLDIASGTAVRFPFG	179
Qy	181	QTRKVQLIPLGGSKKVIQMGNLVNNIADERHKKHAKLDKAKSHGF	224
Db	180	EKSVELIDIGNRRIFGFNALVDRQADNESKKTALHRAKERGF	223

Search completed: August 9, 2006, 21:23:50
Job time : 198 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:29:12 ; Search time 50 Seconds
(without alignments)
395.638 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLLYAGEVA.....ADERHKHAKDKAKSHGFIK 226

Scoring table: BLOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents_AA.*
- 1: /EMC Celerra_SIDS3/prodata/2/iaa/5 COMB.pep.*
 - 2: /EMC Celerra_SIDS3/prodata/2/iaa/6 COMB.pep.*
 - 3: /EMC Celerra_SIDS3/prodata/2/iaa/7 COMB.pep.*
 - 4: /EMC Celerra_SIDS3/prodata/2/iaa/H COMB.pep.*
 - 5: /EMC Celerra_SIDS3/prodata/2/iaa/PCTUS COMB.pep.*
 - 6: /EMC Celerra_SIDS3/prodata/2/iaa/RE COMB.pep.*
 - 7: /EMC Celerra_SIDS3/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	638	54.0	234	2	US-09-338-920B-11
2	616	52.2	238	1	US-08-920-095-2
3	616	52.2	238	1	US-09-338-920B-5
4	616	52.2	238	5	PCT-US96-05800-2
5	616	52.2	245	2	US-09-431-705-4
6	611	51.7	237	1	US-08-467-822-22
7	611	51.7	237	1	US-08-432-697-22
8	611	51.7	237	2	US-08-466-248-22
9	609	51.6	238	2	US-09-338-920B-7
10	579	49.0	228	2	US-10-012-819-154
11	578	48.9	237	1	US-08-467-822-20
12	578	48.9	237	2	US-08-432-697-20
13	578	48.9	237	2	US-08-466-248-20
14	572	48.4	237	2	US-09-338-920B-9
15	554	46.9	213	2	US-09-338-920B-13
16	451	38.2	840	1	US-08-467-822-25
17	451	38.2	840	2	US-08-432-697-25
18	451	38.2	840	2	US-08-466-248-25
19	287	24.3	100	1	US-08-967-513-3
20	287	24.3	100	1	US-08-687-645B-3
21	287	24.3	103	2	US-09-489-039A-9261
22	275	23.3	99	1	US-07-732-242C-1
23	273	23.1	145	2	US-09-134-001C-4982
24	272	23.0	107	2	US-09-543-681A-6101
25	270	22.9	132	2	US-09-134-001C-5049
26	268	22.7	103	2	US-09-338-352-5887

27	267	22.6	100	1	US-08-467-822-23	Sequence 23, Appl
28	267	22.6	100	2	US-08-432-697-23	Sequence 23, Appl
29	267	22.6	100	2	US-08-466-248-23	Sequence 23, Appl
30	259	21.9	125	2	US-09-252-991A-26884	Sequence 26884, A
31	256	21.7	100	2	US-09-602-777A-18	Sequence 18, Appl
32	250	21.2	111	2	US-09-710-279-116	Sequence 116, Appl
33	243	20.6	153	2	US-09-543-681A-6085	Sequence 6085, Ap
34	234	19.8	106	1	US-07-732-242C-2	Sequence 2, Appli
35	229	19.4	119	2	US-09-328-352-5915	Sequence 5915, Ap
36	221	18.7	109	1	US-08-467-822-24	Sequence 24, Appl
37	221	18.7	109	2	US-08-432-697-24	Sequence 24, Appl
38	221	18.7	109	2	US-08-466-248-24	Sequence 24, Appl
39	219	18.5	106	1	US-08-967-513-4	Sequence 4, Appli
40	219	18.5	106	1	US-08-687-645B-4	Sequence 4, Appli
41	218	18.5	109	2	US-09-489-039A-9266	Sequence 9266, Ap
42	217	18.4	162	2	US-09-602-777A-10	Sequence 10, Appl
43	197	16.7	137	2	US-09-252-991A-26886	Sequence 26886, A
44	171	14.5	289	2	US-09-134-001C-4415	Sequence 4415, Ap
45	167	14.1	791	2	US-10-104-047-2307	Sequence 2307, Ap

ALIGNMENTS

RESULT 1
US-09-338-920B-11
; Sequence 11, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Helicobacter heilmannii
US-09-338-920B-11

Query Match	54.0%	Score 638;	DB 2;	Length 234;
Best Local Similarity	54.8%;	Pred. No. 6.9e-23;		
Matches 126;	Conservative 41;	Mismatches 55;	Indels 8;	Gaps 5;
Qy	1	VKLTPKEQKFLLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAOLMEE	60	
Db	1	MKLTPKELDKMLHLYAGELAKQKAKGKLNTEAVALISAHVMEARAGKKSVDLMQE	60	
Qy	61	CMHFLKKDEVNPGVGNMVPDLGVEATFPDGTGLVTNNWPIEP-DEHFKAGEV--KFGCDK	117	
Db	61	GRTLLKADDVMPGVAHMIHEVGIEAGFPDGTGLVTIHTVPEAGSDKLAPEVILK--NE	117	
Qy	118	DIENAGKEVTELEVNTNEGPKSLHVSHFHPFEAKALKFDPREKAYGKRLDIPSGNTLRI	177	
Db	118	DIENAGKHAQLKVKNNKGDPRVQVQSHFHPFEVNNKLLDFDREKAYGKRLDIASGTAVER	177	
Qy	178	GAGQTRKVLPLGSGKKVIGNNGIANNIADERHKHAKLDKAK-SH-GFI	225	
Db	178	EPGBEKTVLLDIGNKKRIYGNALVDROADHGGKLLAKRAKKEHFGTI	227	

RESULT 2
US-08-920-095-2
; Sequence 2, Application US/08920095
; Patent No. 5837240
; GENERAL INFORMATION:
; APPLICANT: Cynthia K. Lee et al.

```
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,095
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,041
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-920-095-2

Query Match      52.2%; Score 616; DB 1; Length 238;
Best Local Similarity 52.7%; Pred. No. 8.6e-22;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

Qy      1 VKLTPKEQKFLYYAGVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db      1 MKLTPKELDKLMLHYAGELAKKKEGKIKLNYEVALISAHIMEARAGKKTAAELMQE 60
Qy      61 CMHFLKDKDEVNMPGVGNVDPDLGVEATFPDGTGKLVTVNWPIDPEHFHFKAGEVKFGCDKIE 120
Db      61 GRTLKPDVMDGVASMIHEVGIEAMFPDGTGKLVTVHTPIEANGKLVPEGL-FLKNEDIT 119
Qy      121 LNAGKEVTELEVTNEGPKSLHVGSHFHPFEANKALKPDREKAYGKRLDIPSGNTLIRIGAG 180
Db      120 INEGKAVSVKVNQVGNDRPVQIGSHFHFPEVNRCLDFDREKTFGKRLDIASGTAVRPEFG 179
Qy      181 QTRKVQLIPLGSKKVGNGVLNNIADERHKHAKDIAKSHGF 224
Db      180 BEKSVELIDIGNRRIFGNALVDRQADNESKIALHRAKERGF 223

; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05800
; FILING DATE: 23-APR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,041
; FILING DATE: 28-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/568,122
; FILING DATE: 06-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/020001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-05800-2
```

```
; SEQ ID NO 5
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-338-920B-5

Query Match      52.2%; Score 616; DB 2; Length 238;
Best Local Similarity 52.7%; Pred. No. 8.6e-22;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

Qy      1 VKLTPKEQKFLYYAGVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db      1 MKLTPKELDKLMLHYAGELAKKKEGKIKLNTVEVALISAHIMEARAGKKTAAELMQE 60
Qy      61 CMHFLKDKDEVNMPGVGNVDPDLGVEATFPDGTGKLVTVNWPIDPEHFHFKAGEVKFGCDKIE 120
Db      61 GRTLKPDVMDGVASMIHEVGIEAMFPDGTGKLVTVHTPIEANGKLVPEGL-FLKNEDIT 119
Qy      121 LNAGKEVTELEVTNEGPKSLHVGSHFHPFEANKALKPDREKAYGKRLDIPSGNTLIRIGAG 180
Db      120 INEGKAVSVKVNQVGNDRPVQIGSHFHFPEVNRCLDFDREKTFGKRLDIASGTAVRPEFG 179
Qy      181 QTRKVQLIPLGSKKVGNGVLNNIADERHKHAKDIAKSHGF 224
Db      180 BEKSVELIDIGNRRIFGNALVDRQADNESKIALHRAKERGF 223

; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; GENERAL INFORMATION:
; APPLICANT: Oravax, Inc.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05800
; FILING DATE: 23-APR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,041
; FILING DATE: 28-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/568,122
; FILING DATE: 06-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/020001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-05800-2
```

db 179 EKSVELIDIGGNRIFGFNALVDRQADNESKIALHRAKRGF 222

RESULT 7
US-08-432-697-22
; Sequence 22, Application US/08432697
; Patent No. 6248330

```

: TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
: TITLE OF INVENTION: HELICOBLACTER INFECTION, POLYPEPTIDES FOR USE IN THE
: TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
: TITLE OF INVENTION: POLYPEPTIDES
: NUMBER OF SEQUENCES: 44
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flanagan, Henderson, Farabow, Garrett &
: ADDRESS: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US 08/466,248
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/447,177
: FILING DATE: 19-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/432,697
: FILING DATE: 02-MAY-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Meyers, Kenneth J.
: REGISTRATION NUMBER: 25,146
: REFERENCE/DOCKET NUMBER: 03495.0137-02000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 408-4000
: TELEFAX: (202) 408-4400
: INFORMATION FOR SEQ ID NO: 22:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 237 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-466-248-22

Query Match 51.7%; Score 611; DB 2; Length 237;
Best Local Similarity 52.7%; Pred. No. 1.5e-21;
Matches 118; Conservative 43; Mismatches 61; Indels 2; Gaps 2;

Qy 1 VKLTPKQSEKFLLYAGVAVRKRAEGLKLNQPEAIYATSAHIMDEARRKKTVAQLMEE 60
Db 1 MKLTPKELDK-LMHYAGELAKRKEKIGIKLNYVEAVALISAHIMEERAGKKTAAELMQE 59
Qy 61 CMHFELKDEVMPPGVGNVPYDILGVEATFPDGTKLVTNWPTEPDHFHFKAGVKGCDKDIE 120
Db 60 GRTLLKPDVNDMGVASMTHIEVGIEAMFPDGTKLVTVHTPIEANGKLVPGL-FLKNEDIT 118
Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHPFFPEANKALKFDRKAYGKRKLDIPSGNLTIRICAG 180
Db 119 INEGKAVSVKXNVGDRPVQIGSHFHFPEVNRCLDFDREKTFGKRLLDIASGTAVRPEG 178
Qy 181 QTRKVQLPIPLGSKSKVIGNGLVNNIADRRHKHKAIDKAKSHGF 224
Db 179 EEKSVELIDIGNRRIFGFLNALVDRQADNESKKIALHRAKERGF 222

```

RESULT 9
US-09-338-920B-7
; Sequence 7, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan

```

; MOLECULE TYPE: protein
US-08-466-248-22

; COLOR: 1.1near.

Query Match          51.7%;   Score 611;   DB 2;   Length 237;
Best Local Similarity 52.7%;   Pred. No. 1.5e-21;
Matches 118;   Conservative 43;   Mismatches 61;   Indels 2;   Gaps 2;

Qy 1 VKLTPKQEKFLYYAGEVARKKASGLKLNQPEAIAYISAHIMDEARRCKTKTVAOLMEE 60
    :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 1 MKLTPKELDK-LMHYAGELAKKKEKGKIKLYIVAEVALISAHIMDEARRAGAKTAAELMQE 59
    :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

Qy 61 CMHFLKQDSVMPGVGNMVDLGVATEFPDGTKLVTWNPLEPDHFKAGKVGKCDKIE 120
    :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 60 GRTLKPDVMDGVAGMIEHVGEIAMPDGTKLVTVHTPIEANGKLVPGSL-FLKNEDIT 118
    :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

Qy 121 LNAGKEVTELEVNTNEGPKSLIHVSGSHFFHFEANKALKPDREKAYGKKRLDIPSGNLTIRIGAG 180
    :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 119 INEGKKAIVSVKVNVDGRPVQIGSHFFHFFVFNCLDPDREKTFGKRLDIASGTAVRPEFG 178
    :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

Qy 181 QTRKVLQILPGGSKKVIYGMNGLVNNITADERHKKHALDKAKSHGF 224
    :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 179 EEKSVELIDIGGNRRIFGFNALVDROADNESKKIALHRAKERGF 222
    :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

```

```
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-338-920B-7

Query Match          51.6%; Score 609; DB 2; Length 238;
Best Local Similarity 52.2%; Pred. No. 1.9e-21;
Matches 11; Conservative 43; Mismatches 63; Indels 1; Gaps 1;

Qy   1 VKLTPKEQKFLYYAGEVARKKAEGKLKNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db   1 MKLTPKELDKLMHLYAGELAKRKEKGKLNLYEAVRLISAHIEMEARRGKKTAAELMQE 60

Qy   61 CMHFPLKDFVMPPGVGNMVPDLGEATPDPDGTKLVTVNWPIEDPDHFHKAGEVKFCKDIE 120
      |||:||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db   61 GRTLKKDDVDMDGVASMIHEVGIEAMPDPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119

Qy   121 LNAKEVTELEVTNEGPKSLHVSHGFHFFFEANKALKFDREKAYGKRLLDIPSGNTLRIGAG 180
      :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db   120 INEGKKAIVSKVPVGPDRPVQIGHFHFFEVNRCCLDFDREKTFGKRLLDIASGTAVRPFG 179

Qy   181 QTRKVQLPIGGSSKVIYGMNLVNNIADERHKHKALDKAKSHGF 224
      :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db   180 BEKSVELIDIGNERRIFGFNALVDROQADNESKIALHRAKERGF 223

RESULT 10
US-10-012-819-154
; Sequence 154, Application US/10012819
; Patent No. 6916615
; GENERAL INFORMATION:
; APPLICANT: Legrain, Pierre
; APPLICANT: Selig, Luc
; APPLICANT: Rain, Jean-Christophe
; TITLE OF INVENTION: Collection of prokaryotic DNA for Two-Hybrid System, Helicobacter
; TITLE OF INVENTION: pylori Protein-Protein Interactions and Applications thereof
; FILE REFERENCE: B5053
; CURRENT APPLICATION NUMBER: US/10/012,819
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: EP 99401066.8
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 154
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-10-012-819-154

Query Match          49.0%; Score 579; DB 2; Length 228;
Best Local Similarity 52.1%; Pred. No. 5e-20;
Matches 11; Conservative 41; Mismatches 60; Indels 1; Gaps 1;

Qy   12 LLYYAGEVARKKAEGKLKNQPEAIAYISAHIMDEARRGKKTVAQLMECHFLKDEVM 71
      :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db   2 MLHYAGELAKRKEKGKLNLYEAVRLISAHIEMEARAGKTAELMQEGRITLLKPDDVM 61

Qy   72 PGVGNMVPDLGEATPDPDGTKLVTVNWPIEDPDHFHKAGEVKFCKDIEIENAGKEVTELE 131
      :||:|:|:|:|:|:~::~||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db   62 DGVASMIHEVGIEAMPDPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDITNEGKKAIVSK 120

Qy   132 VTNNEGPKSLHVSHGFHFFFEANKALKFDREKAYGKRLLDIPSGNTLRIGAGQTRKVQLPIGL 191
```

Db 1 MKLTPKELDKMLHYAGRLAERLARGVKLVNTEAVALISGRVMEKARDGNKSVADLMQE 60
 Qy 61 CMHFLKDEVMGPNVDPDLGVEATFPDGTGLVTNNPPIBPDEHFKAAGEVKGCDKDIE 120
 Db 61 GRTWLKKNVMDGVASMIHEVGIEANFPDGTGLVTIHTPVEDNGKLAPGEV-FLKNEDIT 119
 Qy 121 LNAKEVTEVLTNEGPKSLVHGVSHFHFPEANKALKFDEKAYGKRLDIPSGNTLRIGAG 180
 Db 120 INAGKEAISLVKKNKGRDPVQVGVSHFHFPEVFNKLLDFDRAKSFCKRLDIASGTAVRPEFG 179
 Qy 181 OTRKVLPIPLGSGKKVIGMGNLVNNIADERHKKHAKDKAKSHGF 224
 Db 180 EKSVELIDIGGNKRIYGFNSLVDRQADAGKUGLKRKAKGKF 223

RESULT 12
 US-08-432-697-20
 ; Sequence 20, Application US/08432697
 ; Patent No. 6248330
 ; GENERAL INFORMATION:
 ; APPLICANT: Labigne, Agnes
 ; APPLICANT: Sauerbaum, Sebastien
 ; APPLICANT: Ferrero, Richard L.
 ; APPLICANT: Thiberge, Jean-Michel
 ; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
 ; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
 ; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
 ; TITLE OF INVENTION: POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 44
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 ; ADDRESSEE: Dunner
 ; STREET: 1300 I Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/432,697
 ; FILING DATE: 02-MAY-1995
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meyers, Kenneth J.
 ; REGISTRATION NUMBER: 25,146
 ; REFERENCE/DOCKET NUMBER: 03495.0137-00000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 408-4000
 ; TELEFAX: (202) 408-4400
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 237 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FEATURE:
 ; NAME/KEY: Protein
 ; LOCATION: 1-237
 ; OTHER INFORMATION: /note= "URE A - FIGURE 3."
 ; US-08-432-697-20

Query Match 48.9%; Score 578; DB 2; Length 237;
 Best Local Similarity 50.4%; Pred. No. 6.2e-20;
 Matches 113; Conservative 43; Mismatches 67; Indels 1; Gaps 1;
 Qy 1 VKLTPKEQKFLYYAGVARRKRAEGLKNOPEAIYISAHIMDEARRGKKTVAQLME 60
 Db 1 MKLTPKELDKMLHYAGRLAERLARGVKLVNTEAVALISGRVMEKARDGNKSVADLMQE 60

Qy 61 CMHFLKDEVMGPNVDPDLGVEATFPDGTGLVTNNPPIBPDEHFKAAGEVKGCDKDIE 120
 Db 61 GRTWLKKNVMDGVASMIHEVGIEANFPDGTGLVTIHTPVEDNGKLAPGEV-FLKNEDIT 119
 Qy 121 LNAKEVTEVLTNEGPKSLVHGVSHFHFPEANKALKFDEKAYGKRLDIPSGNTLRIGAG 180
 Db 120 INAGKEAISLVKKNKGRDPVQVGVSHFHFPEVFNKLLDFDRAKSFCKRLDIASGTAVRPEFG 179
 Qy 181 OTRKVLPIPLGSGKKVIGMGNLVNNIADERHKKHAKDKAKSHGF 224
 Db 180 EKSVELIDIGGNKRIYGFNSLVDRQADAGKUGLKRKAKGKF 223

RESULT 13
 US-08-466-248-20
 ; Sequence 20, Application US/08466248
 ; Patent No. 6258359
 ; GENERAL INFORMATION:
 ; APPLICANT: Labigne, Agnes
 ; APPLICANT: Sauerbaum, Sebastien
 ; APPLICANT: Ferrero, Richard L.
 ; APPLICANT: Thiberge, Jean-Michel
 ; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
 ; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
 ; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
 ; TITLE OF INVENTION: POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 44
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 ; ADDRESSEE: Dunner
 ; STREET: 1300 I Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/466,248
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/447,177
 ; FILING DATE: 19-MAY-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/432,697
 ; FILING DATE: 02-MAY-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meyers, Kenneth J.
 ; REGISTRATION NUMBER: 25,146
 ; REFERENCE/DOCKET NUMBER: 03495.0137-02000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 408-4000
 ; TELEFAX: (202) 408-4400
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 237 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FEATURE:
 ; NAME/KEY: Protein
 ; LOCATION: 1-237
 ; OTHER INFORMATION: /note= "URE A - FIGURE 3."
 ; US-08-466-248-20

Query Match 48.9%; Score 578; DB 2; Length 237;
Best Local Similarity 50.4%; Pred. No. 6.2e-20;
Matches 113; Conservative 43; Mismatches 67; Indels 1; Gaps 1;
QY 1 VKLTPKEOEKPLLYAGVAVKRRKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
DB 1 MKLTPKELDKMLHYAGRLAEERLARGVKLNYTEAVALISGRVMEKARDGKNSVADLMQE 60
QY 61 CMHFLKDEVMGPNVGNMVPDLGVEATPDGKLTVTNNPIEPDEHFHFKAGEVKGCDKDIE 120
DB 61 GRTWLKKNVMDGVASMIHEVGIEANFPDGTGLVTIHTPVEDNGKLAPGEV-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180
DB 120 INAGKEAISLKVKNKGRDPVQVGHFHFPEVKNKLLDFDRAKSFCKRLDIASGTAVRPEFG 179
QY 181 QTRKVLQIPLGSKKVIKGMNGLVNNIADERRHKHAKLDKAKSHGF 224
DB 180 EKSVELIDIGGNKRIYGFNSLVRQADADGKGLGKRAKEKGF 223
RESULT 14
US-09-338-920B-9
; Sequence 9, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-338-920B-9

Query Match 48.4%; Score 572; DB 2; Length 237;
Best Local Similarity 50.0%; Pred. No. 1.2e-19;
Matches 112; Conservative 43; Mismatches 68; Indels 1; Gaps 1;
QY 1 VKLTPKEOEKPLLYAGVAVKRRKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
DB 1 MKLTPKELDKMLHYAGRLAEERLARGVKLNYTEAVALISGRVMEKARDGKNSVADLMQE 60
QY 61 CMHFLKDEVMGPNVGNMVPDLGVEATPDGKLTVTNNPIEPDEHFHFKAGEVKGCDKDIE 120
DB 61 GRTWLKKNVMDGVASMIHEVGIEANFPDGTGLVTIHTPVEDNGKLAPGEV-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180
DB 120 INAGKEAISLKVKNKGRDPVQVGHFHFPEVKNKLLDFDRAKSFCKRLDIASGTAVRPEFG 179
QY 181 QTRKVLQIPLGSKKVIKGMNGLVNNIADERRHKHAKLDKAKSHGF 224
DB 180 EKSVELIDIGGNKRIYGFNSLVRQADADGKGLGKRAKEKGF 223

RESULT 15
US-09-338-920B-13
; Sequence 13, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease

FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Helicobacter mustelae
US-09-338-920B-13
Query Match 46.9%; Score 554; DB 2; Length 213;
Best Local Similarity 52.3%; Pred. No. 7.1e-19;
Matches 113; Conservative 37; Mismatches 61; Indels 5; Gaps 2;
QY 13 LYVAGEVAVKRRKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEECMHFLKDEVM 72
DB 1 LHYAGELAKKRRKRGVKNLYTEAVALISMEIHAREGKKTVAQLMEECMHFLKDEVM 60
QY 73 GVGNNMVPDLGVEATPDGKLTVTNNPIEPDEHFHFKAGE--VKFGCDKDIELNAGKEVTEL 130
DB 61 GVAALVHEVQIEVCFPDGTLTLLVTVHNPENNGKLFHGFILK--DEDIVLNAGKEAIEV 117
QY 131 EVTNEGPKSLHVGSHFHFPEANKALKPDREKAYGKRLDIPSGNTLRIGAGQTRKVLQIPL 190
DB 118 KVSNGKDRPIQVGHFHFPEANKALKPDREKAYGKRLDIPSGNTLRIGAGQTRKVLQIPL 177
QY 191 GSKKVIKGMNGLVNNIADERRHKHAKLDKAKSHGF 226
DB 178 GGNQRIFGFNDLNGQVNEEDNKRKALAAKAKGFIK 213
Search completed: August 9, 2006, 21:30:35
Job time : 51 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:29:51 ; Search time 184 Seconds
(without alignments)
568.949 Million cell updates/sec

Title: US-09-904-994B-2
Perfect score: 1181
Sequence: 1 VKLTPKQEKFLYYAGEVA.....ADERHKHALDKAKSHGFIK 226

Scoring table: BLOSUM62
Gapop 4.0 , Gapext 1.0

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main.*
1: /EMC_Celerra_SID33/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /EMC_Celerra_SID33/prodata/2/pubpaa/US08_PUBCOMB.pep.*
3: /EMC_Celerra_SID33/prodata/2/pubpaa/US09_PUBCOMB.pep.*
4: /EMC_Celerra_SID33/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
5: /EMC_Celerra_SID33/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
6: /EMC_Celerra_SID33/prodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1181	100.0	226	3	US-09-904-994B-2
2	1178	99.7	226	3	US-09-904-994B-14
3	1174	99.4	226	3	US-09-904-994B-5
4	1174	99.4	226	3	US-09-904-994B-11
5	1170	99.1	226	3	US-09-904-994B-8
6	665	56.3	234	4	US-10-639-273-3
7	638	54.0	234	4	US-10-639-273-34
8	619	52.4	238	3	US-09-815-242-11468
9	616	52.4	238	3	US-10-335-977-8629
10	605	51.2	237	4	US-09-815-242-11286
11	582	49.3	228	4	US-10-639-273-36
12	579	49.0	228	4	US-10-335-977-8628
13	572	48.4	227	4	US-10-012-819-154
14	556	47.1	225	4	US-10-639-273-35
15	530	44.9	811	4	US-10-639-273-37
16	516	43.7	207	4	US-10-282-122A-70574
17	509	43.1	226	3	US-10-282-122A-70447
18	505	42.8	779	4	US-09-815-242-5448
19	493	41.7	207	4	US-10-282-122A-69134
20	492	41.7	207	4	US-10-282-122A-45038
21	490	41.5	261	4	US-10-282-122A-48209
22	487	41.2	209	4	US-10-282-122A-55507
23	487	39.5	228	4	US-10-282-122A-68251
24	467	39.1	840	4	US-10-282-122A-50077
25	462	39.1	840	4	US-10-621-833-7
26	462	39.1	840	5	US-10-731-877-1
27	462	39.1	840	6	US-11-046-271-1

28	436	36.9	843	4	US-10-437-963-124057
29	430	36.4	171	4	US-10-335-977-8627
30	415	35.1	837	4	US-10-424-599-254635
31	408	34.5	227	4	US-10-156-761-10253
32	382	32.3	837	4	US-10-418-962-2
33	317	26.8	227	4	US-10-425-115-201340
34	291	24.6	121	4	US-10-282-122A-49772
35	287	24.3	100	3	US-09-815-242-11745
36	287	24.3	100	4	US-10-156-761-14633
37	287	24.3	100	4	US-10-282-122A-60250
38	284	24.0	100	4	US-10-282-122A-62539
39	284	24.0	100	4	US-10-282-122A-64616
40	283	24.0	100	4	US-10-282-122A-56822
41	283	24.0	100	4	US-10-282-122A-56846
42	276	23.4	100	4	US-10-282-122A-59398
43	273	23.1	100	3	US-09-815-242-11047
44	273	23.1	100	4	US-10-282-122A-58211
45	273	23.1	145	4	US-10-724-972A-6468

ALIGNMENTS

RESULT 1

US-09-904-994B-2
; Sequence 2, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-2

Query Match 100.0%; Score 1181; DB 3; Length 226;
Best Local Similarity 100.0%; Pred. No. 3.3e-42;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VKLTPKQEKFLYYAGEVARRKKAEGKLNOPEAIAYISAHIMDEARRGKKTVAOLMEE 60
Db 1 VKLTPKQEKFLYYAGEVARRKKAEGKLNOPEAIAYISAHIMDEARRGKKTVAOLMEE 60
Qy 61 CMHFLKQDEVMPGVGNMVPDLGVEATFPDGTKLVTNNWPIEDPEHFKAGEVKGCDKDIE 120
Db 61 CMHFLKQDEVMPGVGNMVPDLGVEATFPDGTKLVTNNWPIEDPEHFKAGEVKGCDKDIE 120
Qy 121 LNAGKEVTELVNTEGPKSLHVGSHFHPFEANKALKFDEKAYGKRLDIPSGNTLIRIGAG 180
Db 121 LNAGKEVTELVNTEGPKSLHVGSHFHPFEANKALKFDEKAYGKRLDIPSGNTLIRIGAG 180
Qy 181 QTRKVLPLGSGKKVIGMGNLVNNIADERHKHALDKAKSHGFIK 226
Db 181 QTRKVLPLGSGKKVIGMGNLVNNIADERHKHALDKAKSHGFIK 226

RESULT 2

US-09-904-994B-14
; Sequence 14, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21


```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-14

Query Match      99.7%; Score 1178; DB 3; Length 226;
Best Local Similarity 99.6%; Pred. No. 4.4e-42;
Matches 225; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGVARRKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGVARRKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60

Qy 61 CMHFLKKDDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Db 61 CMHFLKKDDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGAG 180

Qy 181 QTRKVQLIPLGGSKKVGNGLVNNIADERRHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVGNGLVNNIADERRHKHKKALDKAKSHGFIK 226

RESULT 3
US-09-904-994B-5
; Sequence 5, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-5

Query Match      99.4%; Score 1174; DB 3; Length 226;
Best Local Similarity 99.1%; Pred. No. 6.6e-42;
Matches 224; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGVARRKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGVARRKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60

Qy 61 CMHFLKKDDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Db 61 CMHFLKKDDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGAG 180

Qy 181 QTRKVQLIPLGGSKKVGNGLVNNIADERRHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVGNGLVNNIADERRHKHKKALDKAKSHGFIK 226

RESULT 4
US-09-904-994B-11
; Sequence 11, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
```

```
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-11

Query Match      99.4%; Score 1174; DB 3; Length 226;
Best Local Similarity 99.1%; Pred. No. 6.6e-42;
Matches 224; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGVARRKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGVARRKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60

Qy 61 CMHFLKKDDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Db 61 CMHFLKKDDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGAG 180

Qy 181 QTRKVQLIPLGGSKKVGNGLVNNIADERRHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVGNGLVNNIADERRHKHKKALDKAKSHGFIK 226

RESULT 5
US-09-904-994B-8
; Sequence 8, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-8

Query Match      99.1%; Score 1170; DB 3; Length 226;
Best Local Similarity 98.7%; Pred. No. 9.8e-42;
Matches 223; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGVARRKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGVARRKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60

Qy 61 CMHFLKKDDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Db 61 CMHFLKKDDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGAG 180

Qy 181 QTRKVQLIPLGGSKKVGNGLVNNIADERRHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVGNGLVNNIADERRHKHKKALDKAKSHGFIK 226

RESULT 6
US-09-904-994B-8
; Sequence 8, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-8

Query Match      99.1%; Score 1170; DB 3; Length 226;
Best Local Similarity 98.7%; Pred. No. 9.8e-42;
Matches 223; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGVARRKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGVARRKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60

Qy 61 CMHFLKKDDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Db 61 CMHFLKKDDEVMPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGAG 180

Qy 181 QTRKVQLIPLGGSKKVGNGLVNNIADERRHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVGNGLVNNIADERRHKHKKALDKAKSHGFIK 226
```

```
US-10-639-273-3
; Sequence 3, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Helicobacter bizzozeronii
US-10-639-273-3

Query Match      56.3%; Score 665; DB 4; Length 234;
Best Local Similarity 57.3%; Pred. No. 5.6e-20;
Matches 130; Conservative 39; Mismatches 52; Indels 6; Gaps 3;

Qy 1 VKLTPKEQKFLYYAGEVARKRAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRANGKLVNTEAVALLISAHVMEERAGKKSVDLMOE 60

Qy 61 CMHFLKDEVMGPNVDPDLGVEATPPDGTGLVTVNWPIDPEHFKA-GEV--KFGCDK 117
Db 61 GRTLLKADDDVMPGVAMHIEVGIEANFPDGTGLVTIHTPVEDGGHKLAPGSEVILK--NE 117

Qy 118 DIENAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKDFREKAYGKRLDIPSGNTLRI 177
Db 118 DITLNAGKQATTLTEVHNKGRDPVQVGSHPHFVFNKLLDFPREKAYGKRLDIASGTAVRF 177

Qy 178 GAGOTRQVQLIPLGGSKKVIKGMNGLVNNIADERHKHAKLDKAK-SH-GFI 225
Db 178 EPGEKTVELIDIGGNRIYGFNSLVDRQADTDGKKLAKRAKEHGF 224

US-10-639-273-34
; Sequence 34, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Helicobacter heilmannii
US-10-639-273-34

Query Match      54.0%; Score 638; DB 4; Length 234;
Best Local Similarity 54.8%; Pred. No. 8.1e-19;
Matches 126; Conservative 41; Mismatches 55; Indels 8; Gaps 5;

Qy 1 VKLTPKEQKFLYYAGEVARKRAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKQKAGIKLNYTEAVALLISAHVMEERAGKKSVDLMOE 60
```

```
Qy 61 CMHFLKDEVMGPNVDPDLGVEATPPDGTGLVTVNWPIDPEHFKA-GEV--KFGCDK 117
Db 61 GRTLLKADDDVMPGVAMHIEVGIEANFPDGTGLVTIHTPVEAGSDKLAPGSEVILK--NE 117

Qy 118 DIENAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKDFREKAYGKRLDIPSGNTLRI 177
Db 118 DITLNAGKQATTLTEVHNKGRDPVQVGSHPHFVFNKLLDFPREKAYGKRLDIASGTAVRF 177

Qy 178 GAGOTRQVQLIPLGGSKKVIKGMNGLVNNIADERHKHAKLDKAK-SH-GFI 225
Db 178 EPGEKTVELIDIGGNRIYGFNSLVDRQADTDGKKLAKRAKEHGF 224

RESULT 8
US-09-815-242-11468
; Sequence 11468, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11468
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11468

Query Match      52.4%; Score 619; DB 3; Length 238;
Best Local Similarity 53.1%; Pred. No. 5.6e-18;
Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;
```

```
Qy 1 VKLTPKEQKFLYYAGEVARKRAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELARKRKKGKIKLNTVEAVALLISAHIMEERAGKKTAAELMOE 60

Qy 61 CMHFLKDEVMGPNVDPDLGVEATPPDGTGLVTVNWPIDPEHFKA-GEVFGCDKDIE 120
Db 61 GRTLLKPDVDMGVASMIHEVGIEANFPDGTGLVTIHTPIEANGKLVPGEL-FLKNEDIT 119

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKDFREKAYGKRLDIPSGNTLRI 180
Db 120 INEGKAVSVKVNKVDPRVQVIGSHFHFPEVNRCLDFREKTFGKRLDIASGTAVRFEG 179

Qy 181 QTRKVLQIPLGGSKKVIKGMNGLVNNIADERHKHAKLDKAKSHGF 224
Db 180 BEKSVELIDIGGNRIYGFNSLVDRQADNESKKIALHRAKERGF 223
```

RESULT 9
 US-10-335-977-8629
 ; Sequence 8629, Application US/10335977
 ; Publication No. US20040052799A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DOUGLAS SMITH et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 ; RELATING TO HELICOBACTER PYLORI FOR
 ; DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 10031
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 28 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109-1875
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: Windows NT 4.0
 ; SOFTWARE: UNIX
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/335,977
 ; FILING DATE: 30-Dec-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/993,002
 ; FILING DATE: 17-DEC-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mandragoras, Amy E.
 ; REGISTRATION NUMBER: 36,207
 ; REFERENCE/DOCKET NUMBER: GTN-018
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)742-4214
 ; INFORMATION FOR SEQ ID NO: 8629:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 238 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHEICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Helicobacter pylori
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (B) LOCATION 1...238
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 8629:
 US-10-335-977-8629
 Query Match 52.4%; Score 619; DB 4; Length 238;
 Best Local Similarity 53.1%; Pred. No. 5.6e-18;
 Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;
 QY 1 VKLTPKEQKFLYYAGVARKKAEGLKNOPEALAYISAHIMDEARRGKKTVAOLMEE 60
 Db 1 MKLTPKELDKMLHYAGELARKKEGKIKLNTVEAVALISAHIMEERAGKKTAAELMQE 60
 QY 61 CMHFLKDEVMFPGVGNMVPDLGVEATFPDGTKLVTNNWPIEDPHFKAGVFGCDKQIE 120
 Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
 QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHPFEANKALKFDEKAYGKRLDIPSGNLTIRIGAG 180
 Db 120 INEGKAVSVKVNKNDGPRVQIGSHFHPFEVNRCLDFDREKTFGKRLDIASGTAVRFEFG 179
 QY 181 QTRKVLPIGSKKVIKMGVLNVIADERHKKHAKDKAKSHGF 224
 Db 180 EKSVELIDIGNNRRIFGNALVDROADNESKKIALHRAKERGF 223

RESULT 10
 US-09-815-242-11286
 ; Sequence 11286, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11286
 ; LENGTH: 238
 ; TYPE: PRT
 ; ORGANISM: Helicobacter pylori
 US-09-815-242-11286
 Query Match 52.2%; Score 616; DB 3; Length 238;
 Best Local Similarity 52.7%; Pred. No. 7.5e-18;
 Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;
 QY 1 VKLTPKEQKFLYYAGVARKKAEGLKNOPEALAYISAHIMDEARRGKKTVAOLMEE 60
 Db 1 MKLTPKELDKMLHYAGELARKKEGKIKLNTVEAVALISAHIMEERAGKKTAAELMQE 60
 QY 61 CMHFLKDEVMFPGVGNMVPDLGVEATFPDGTKLVTNNWPIEDPHFKAGVFGCDKQIE 120
 Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
 QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHPFEANKALKFDEKAYGKRLDIPSGNLTIRIGAG 180
 Db 120 INEGKAVSVKVNKNDGPRVQIGSHFHPFEVNRCLDFDREKTFGKRLDIASGTAVRFEFG 179
 QY 181 QTRKVLPIGSKKVIKMGVLNVIADERHKKHAKDKAKSHGF 224
 Db 180 EKSVELIDIGNNRRIFGNALVDROADNESKKIALHRAKERGF 223

RESULT 11
 US-10-639-273-36
 ; Sequence 36, Application US/10639273
 ; Publication No. US20040142343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chang, Yung-Fu
 ; APPLICANT: Simpson, Kenneth W
 ; APPLICANT: Zhu, Jiaqian
 ; TITLE OF INVENTION: HELICOBACTER BIZZOZERONI UREASE GENES AND THEIR USES
 ; FILE REFERENCE: IN DIAGNOSTIC AND TREATMENT METHODS
 ; CURRENT APPLICATION NUMBER: US/10/639,273

RESULT 14

US-10-639-273-35
; Sequence 35, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; PRIOR FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-10-639-273-35

Query Match 48.4%; Score 572; DB 4; Length 237;

Best Local Similarity 50.0%; Pred. No. 5.8e-16;

Matches 112; Conservative 43; Mismatches 68; Indels 1; Gaps 1;

Qy	1	VKLTPKEQKELLYAGVARRKRAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE	60
Db	1	MKLTPKELDKLMLHYAGRLAEALARGVKLNTYEAVALISGRVMEKARDGNKSVADLMQE	60
Qy	61	CMHFLKDEVMVPGVGNMVPDLGVZATFPDGTGLVTNNWPIEPDEHFHFKAGEVKGCDKDIE	120
Db	61	GRTWLKKNVMDGVASMIHEVGIEANFPDGTGLVTIHPVEDNGKLAFGEV-FLKNEDIT	119
Qy	121	LNAGKEVTELVNTEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIG	180
Db	120	INAGKEAISLVKNKGDPRPQVGSHFHFPEVKNLLDFDRAKSFCKRLDIASGTAVRFPFG	179
Qy	181	QTRKVQLIPLGGSKKVTGMNGLVNNIADERHKKALDKAKSHGF	224
Db	180	EKSVELIDIGGNRIYGNFSLNDRQADADGKGLGKRAKEKGF	223

RESULT 15

US-10-639-273-37
; Sequence 37, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Helicobacter hepaticus
US-10-639-273-37

Query Match 47.1%; Score 556; DB 4; Length 225;

Best Local Similarity 49.3%; Pred. No. 2.5e-15;

Matches 113; Conservative 45; Mismatches 62; Indels 9; Gaps 6;

Qy	1	VKLTPKEQKELLYAGVARRKRAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE	60
----	---	---	----

Db	1	MRLTPKELDKLMLHYAGALAKSRKRGTKLNTVESTALISMEIMELAREGNKSVBELMQQ	60
Qy	61	CMHFLKDEVMVPGVGNMVPDLGVZATFPDGTGLVTNNWPIEPDEHFHFKAGE--VKFGCDKD	118
Db	61	GREILKADVMVEGVASWNEVQVVSFPDGTGLVTIHNPIEDNGKLTPEYILK---DED	117
Qy	119	IELNAGKEVTELVNTEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIG	178
Db	118	ITLNANKESISIKVTHKGDPRPQVGSHFHFPEVKNLLDFDRAQAFGKRLDIASTSVRFE	177
Qy	179	AGQTRKVQLIPLGGSKKVTGMNGLVNNIADERHKKAL-DKAKSHGFI	225
Db	178	PGEKQVNLIDFGGKOKIIGFNDLTNAHI-NKENKEOCLANAAQKH-FI	224

Search completed: August 9, 2006, 21:33:46

Job time : 185 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	436	36.9	843	6	US-10-449-902-47195	Sequence 47195, A
2	401	34.0	841	6	US-10-953-349-33825	Sequence 33825, A
3	375	31.8	836	7	US-11-292-431-37	Sequence 37, Appl
4	276	23.4	784	6	US-10-953-349-33826	Sequence 33826, A
5	271	22.9	100	6	US-10-471-571A-3256	Sequence 3256, Ap
6	251	21.3	136	6	US-10-471-571A-3278	Sequence 3278, Ap
7	172	14.6	876	6	US-10-449-902-31438	Sequence 31438, A
8	172	14.6	936	6	US-10-449-902-56037	Sequence 56037, A
9	169	14.3	327	7	US-11-056-355B-60476	Sequence 60476, A
10	169	14.3	336	7	US-11-056-355B-60475	Sequence 60475, A
11	166	14.1	508	6	US-10-953-349-6491	Sequence 6491, Ap
12	166	14.1	508	7	US-11-056-355B-29545	Sequence 29545, A
13	166	14.1	508	7	US-11-056-355B-31135	Sequence 31135, A
14	166	14.1	508	7	US-11-056-355B-40436	Sequence 40436, A
15	166	14.1	508	7	US-11-056-355B-98803	Sequence 98803, A
16	166	14.1	508	7	US-11-056-355B-110042	Sequence 110042, A
17	166	14.1	595	6	US-10-953-349-6490	Sequence 6490, Ap
18	166	14.1	595	7	US-11-056-355B-29544	Sequence 29544, A
19	166	14.1	595	7	US-11-056-355B-33134	Sequence 33134, A
20	166	14.1	595	7	US-11-056-355B-40435	Sequence 40435, A
21	166	14.1	595	7	US-11-056-355B-98802	Sequence 98802, A
22	166	14.1	595	7	US-11-056-355B-110041	Sequence 110041, A
23	166	14.1	627	6	US-10-953-349-6489	Sequence 6489, Ap
24	166	14.1	627	7	US-11-056-355B-29543	Sequence 29543, A
25	166	14.1	627	7	US-11-056-355B-33133	Sequence 33133, A

Best Local Similarity 36.1%; Pred. No. 5.6e-06;
Matches 95; Conservative 42; Mismatches 78; Indels 48; Gaps 14;

Qy 1 VKLTPKEQKFLYYAGEAVARKKAEGKLNQPEAIYISAHIMDEARRGKKTVQAQM-- 58
:
Db 1 MOLVPRIDIKLTISNLGLAQRLARGVRLNHAETALIASN-LQLIRDGNNSVADLMTI 59
:
Qy 59 -EFCMHFLKDEVMVGNNVPDLGVBATPPDGTKLVT-VN-----WP 99
:
Db 60 GKE-M--LGRRHVLPSVVATLKQVQEGTPTGTNLITVNPNVCDDGDLEKALYGSFLP 116
:
Qy 100 IEPDEHF-----K-AGEV-KFGCKDKIELNAGKEVTELEVNEGPKSLHVGSHF 146
:
Db 117 VPKKETPPDDDDDYQPEKPGVAIPLKTSKKIELNAGRNIMLUKVTSRGDRPIQVGSYH 176
:
Qy 147 HFEANAKLFKDREKAYGKRLLDPISGNTTLRIGAGQTRKVQLPLGGSKKVI-GMNLVNN 205
:
Db 177 HFIEVNPQLDFRAKAYGYRLDIPAGTISRPEPATKAIPLEVGK-KRIIRG--G-NH 231
:
Qy 206 IA----DERHHKALKDAKASHGF 224
:
Db 232 IAVQGVDFRRVDIEMRLQAGF 254
: :

RESULT 4

US-10-953-349-33826
; Sequence 33826, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33826
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33826

Query Match 23.4%; Score 276; DB 6; Length 784;
Best Local Similarity 33.3%; Pred. No. 0.039;
Matches 69; Conservative 32; Mismatches 64; Indels 42; Gaps 9;

Qy 44 MDEARRGKKTVAQLMEECMHFLKDEVMVGNNVPDLGVBATPPDGTKLVTNVWPIEPD 103
:
Db 1 MD---LGK---QL-----LGRQVLPAVPLYLLHTVQVEGTFVDGTKLVTVDHPISLD 46
:
Qy 104 E-----H--F---KAGEVKFGCDKDI-----LNAGKEVTELEVNEG 137
:
Db 47 DGNLELAHSGSFLPVGPGE-KFPSSD-DVEEYPGEIHVSRSRIVLMHRRALSLKVVNKAD 104
:
Qy 138 KSLHVGHSHFFFEANKALKFDRKAYGKRLLDPISGNTTLRIGAGQTRKVQLPLGGSKKVI 197
:
Db 105 RPIQIGHSHYFIETNPYLVEDRKAYGWRLNILAGTAVRPPEPGDAKSIVLSIGGHKLIIK 164
:
Qy 198 GMNLVNNIADERHKHALDAKASHGF 224
:
Db 165 GGNGIADGPIDSSSLNVVMQVANSF 191
: :

RESULT 5

US-10-471-571A-3256
; Sequence 3256, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A

```

; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31438
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-31438

Query Match      14.6%; Score 172; DB 6; Length 876;
Best Local Similarity 24.2%; Pred. No. 6.6e+02;
Matches 62; Conservative 45; Mismatches 95; Indels 54; Gaps 22;

Qy      14  YYAGEVARKEKAEGKLKLNQPEAIAYISAH--I-----MDEARRGKKTVAQLMEECWHF 64
Db      516 YFSGSL--KDAARSLGV-CTTLKRICRHGISRWPSEKINKVNSLKKI-QTIVNSVHG 571

Qy      65  LKGD--EYMPGVGNMVP--DLGVEATFP--DGTKLVTYNWPIEDPHFKAGEVKFGC---D 116
Db      572 VDRSLQDYPATGSLVPVVSLEKLTTPSCDGLPTSPVGKTVSENSDLKSEB---GCSLPD 628

Qy      117 KDIELNAGKEVTELEVTNEGPKSLHVS--HFHFPEANKALKFDBREKAYGKGLDIPSG--N 173
Db      629 GSORQSCQLQISDVKKKSNE--DEFHIGSGNSDFYGANATAKSNSEVTQGP-L-CPTGAFS 684

Qy      174 TLRT-GAGQTR-KVQLIPLGGS--KKVIGMNG-----LVNNIADER-HKHKA--- 215
Db      685 ALHLKGTCTNPSSSLRPSSESTRNQIVGRNSPSIQOEDLMLDNHAEKDHHPSTSG 744

Qy      216 -LD-----KAKSHGFIK 226
Db      745 MTDSSSGSASSHPTFK 760

RESULT 8
US-10-449-902-56037
; Sequence 56037, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56037
; LENGTH: 936
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-56037

Query Match      14.6%; Score 172; DB 6; Length 936;
Best Local Similarity 24.2%; Pred. No. 7.7e+02;
Matches 62; Conservative 45; Mismatches 95; Indels 54; Gaps 22;

Qy      14  YYAGEVARKEKAEGKLKLNQPEAIAYISAH--I-----MDEARRGKKTVAQLMEECWHF 64

```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:24:06 ; Search time 40 Seconds
(without alignments)
543.625 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181
Sequence: 1 VKLTPKEQKFLYYAGEVA.....ADERHKHAKDKAKSHGFIK 226

Scoring table:

BIOSUM62
Gapop 4.0 , Gapext 1.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	619	52.4	238	2 B71977	urease alpha chain
2	616	52.2	238	1 URKCAP	urease (EC 3.5.1.5)
3	572	48.4	237	1 S35290	urease (EC 3.5.1.5)
4	511	43.3	228	2 D75586	urease, beta/gamma
5	462	39.1	840	1 URJB	urease (EC 3.5.1.5)
6	427	36.2	838	2 A96699	probable urease F1
7	416	35.2	227	2 T29055	urease (EC 3.5.1.5)
8	412	34.9	835	2 T37939	urease (EC 3.5.1.5)
9	287	24.3	100	1 A36138	urease (EC 3.5.1.5)
10	284	24.0	100	2 H70664	probable urea prot
11	283	24.0	100	2 D85603	probable urease st
12	283	24.0	100	2 B90794	urease gamma subun
13	282	23.9	100	2 S74889	urease (EC 3.5.1.5)
14	280	23.7	105	2 B69729	urease (gamma subu
15	276	23.4	100	2 A36950	urease (EC 3.5.1.5)
16	275	23.3	100	2 S47102	urease (EC 3.5.1.5)
17	274	23.2	137	1 S38484	urease (EC 3.5.1.5)
18	273	23.1	100	2 A64076	urease (EC 3.5.1.5)
19	271	22.9	100	2 C90027	urease gamma subun
20	270	22.9	100	2 B47090	urease (EC 3.5.1.5)
21	268	22.7	100	1 S08478	urease (EC 3.5.1.5)
22	267	22.6	100	1 B43719	urease (EC 3.5.1.5)
23	267	22.6	100	2 B83037	urease gamma subun
24	267	22.6	124	2 C69729	urease (beta subun
25	265	22.4	100	2 AD2264	urease gamma chain
26	263	22.3	100	2 S42602	urease (EC 3.5.1.5)
27	262	22.2	100	2 D83681	urease gamma subun
28	257	21.8	100	2 AE2872	urease gamma subun
29	257	21.8	107	2 F97648	urease gamma chain

30	257	21.8	126	2 S47103	urease (EC 3.5.1.5)
31	257	21.8	164	1 S31418	urease (EC 3.5.1.5)
32	254	21.5	108	2 AE2264	urease beta chain
33	252	21.3	100	2 S38483	urease (EC 3.5.1.5)
34	252	21.3	136	2 D90027	urease beta subuni
35	249	21.1	159	2 AB3333	urease (EC 3.5.1.5)
36	248	21.0	100	2 AI0324	urease (EC 3.5.1.5)
37	243	20.6	101	2 AC2872	urease beta subuni
38	243	20.6	102	2 S31417	urease (EC 3.5.1.5)
39	241	20.4	101	1 S42604	urease (EC 3.5.1.5)
40	238	20.2	100	2 AC3333	urease (EC 3.5.1.5)
41	238	20.2	105	2 S76224	hypothetical prote
42	237	20.1	124	2 H82890	urease complex com
43	236	20.0	101	2 A82891	urease complex com
44	236	20.0	138	2 T50709	urease (EC 3.5.1.5)
45	234	19.8	100	2 T50708	urease (EC 3.5.1.5)

ALIGNMENTS

RESULT 1

B71977

urease alpha chain - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004

C:Accession: B71977

; Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathc

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: B71977

A>Status: preliminary

A:Molecule type: DNA

A:Cross-references: 1-238 <ARN>

A:Experimental source: strain J99

C:Genetics:

A:Gene: urea

C:Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease 12K ch

F:1-101/Domain: urease 11K chain homology <U1>

F:108-204/Domain: urease 12K chain homology <U12>

Query Match 52.4%; Score 619; DB 2; Length 238;
Best Local Similarity 53.1%; Pred. No. 1e-15; Indels 1; Gaps 1;
Matches 119; Conservative 42; Mismatches 62;

Qy	1	VKLTPKEQKFLYYAGEVARKKAEGLKINQPEAIAYISAHIMDEARRGKKTVAQLMEE	60
Db	1	MKLTPKELDKLMLHVAEGLARKEKGIKNVVEAVALISAHIMEERAGKKTAAELMQE	60
Qy	61	CMHFLKDEVMGVMGVPDLGVRATPDGTLVTNNVIEPDHPKAGEVFGCDKDIE	120
Db	61	GRTLKPDVMDGVASMIHEVGIEAMFPDGTGLVTVHTPIEANGKLVPGEL-FLKNEDIT	119
Qy	121	LNAGKEVTELVNTEGPKSLVGHSHFPEANKALKFDEKAYKELDIPSGNTLRIGAG	180
Db	120	INEGKAVSVKVKVKNVGRFPVQIGSHFPEVNRCLDFDEKTFGRKLDIASGTAVRFPFG	179
Qy	181	QTRKVLPIPLGSKKVGVMGNLVNNIADERHKHAKDKAKSHGF	224
Db	180	EKSVELLDIGGNRIIFGNALVDQADNESKIALHRAKERGF	223

RESULT 2

URKCAP

urease (EC 3.5.1.5) 26K chain - Helicobacter pylori (strains 26695 and others)

N:Alternate names: urea amidohydrolase; urease alpha chain

C:Species: Helicobacter pylori

C:Date: 30-Sep-1991 #sequence_revision 15-Aug-1997 #text_change 09-Jul-2004

C:Accession: A38537; A64529; S07884; B35306; A41502; A61371; E49215

F;1-101/Domain: urease 11K chain homology <U1>
F;108-204/Domain: urease 12K chain homology <U12>

Query Match	52.2%;	Score 616;	DB 1;	Length 238;
Best Local Similarity	52.7%;	Pred. No. 1.3e-15;		
Matches 118;	Conservative 43;	Mismatches 62;	Indels 1;	Gaps 1;

QY 1 VKLTPKQEKFLLYYAGEVARKRKAESGLKLNQPEATAYISAHIMDEARRGKKTVAQLMEE 60
Db :|||:||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1 MKLTPKELDKMLHYAGELAKKKRKEGIKLNTVEAVALISAHIMEERARCKTKTAALMQE 60

QY 61 CMHFLKDEVPVGNNVPDLGVEATFPDGTKLVTNVNPLEPDEHEFKAGEVKGCCKDIE 120
Db :|||:||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
61 GRTLLKPDDVMGVASMIHEVGIEAFPDGTKLVTVHTPIEANGKLVPGSEL-FLKNEEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVSGSHFFFEANKALKFDREKAYGKRLLDIPSGNLTIRIGAG 180
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
120 INEGKAAVSQKVXNGDRPVQGSFHFFFEVRWCLDFDREKTFGKRLLDTASGTAVRFEPG 179

QY 181 QTRKVQLIPLGGSKKVIYGMGNLVNNTADRRHKHALDKAKSHGF 224
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
180 EEKSVELIDIGGNRIYFGNALVDROADNESKIALHRAKERGF 223

RESULT 3
S35290
urease (EC 3.5.1.5) 26K chain - Helicobacter felis
N;Alternate names: urease alpha chain
C;Species: Helicobacter felis
C;Date: 31-Dec-1993 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C;Accession: S35290
R;Ferroo, R.L.; Labigne, A.
Mol. Microbiol. 9, 323-333, 1993
A;Title: Cloning, expression and sequencing of Helicobacter felis urease genes.
A;Reference number: S35290; MUID:94018627; PMID:8412683
A;Accession: S35290
A;Molecule type: DNA
A;Residues: 1-237 <FER>
A;Cross-references: UNIPROT:Q08715; UNIPARC:UPI0000137D6C; EMBL:X69080; NID:G39
C;Genetics:
A;Gene: ureA
C;Complex: heterodimer
C;Function:
A;Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide,
C;Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease,
C;Keywords: heterodimer; hydrolase; metalloprotein
F;1-101/Domain: urease 11K chain homology <U1>
F;108-204/Domain: urease 12K chain homology <U12>

Query Match	48.4%;	Score 572;	DB 1;	Length 237;
Best Local Similarity	50.0%;	Pred. No. 5.8e-14;		
Matches 112;	Conservative 43;	Mismatches 68;	Indels 1;	Gaps 1;

QY 1 VKLTPKQEKFLLYYAGEVARKRKAESGLKLNQPEATAYISAHIMDEARRGKKTVAQLMEE 60
Db :|||:||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1 MKLTPKELDKMLHYAGRLAEALARGVKLNTVEAVALISGRVWEKARDGNKSVDLMQE 60

QY 61 CMHFLKDEVPVGNNVPDLGVEATFPDGTKLVTNVNPLEPDEHEFKAGEVKGCCKDIE 120
Db :|||:||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
61 GRTWLKENVMGVASMIHEVGIEAFPDGTKLVTITPVEDYNGKLVAPGV-FLKNEEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVSGSHFFFEANKALKFDREKAYGKRLLDIPSGNLTIRIGAG 180
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
120 INAGKEAISLVQNKGDRPVQVSGSHFFFEVNKLLDFDRAKSFCKRLDIASGTAVRFEPG 179

QY 181 QTRKVQLIPLGGSKKVIYGMGNLVNNTADRRHKHALDKAKSHGF 224
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
180 EEKSVELIDIGGNRIYGFNSLVDRQADGKKGLKRAKERGF 223

RESULT 4
D75586
urease, beta/gamma subunit - Deinococcus radiodurans (strain R1)

RESULT 4
D7586
urease, beta/gamma subunit - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 06-Jan-2003
C:Accession: D75586
R:White, O.; Easen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: D75586
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-228 <WHI>
A:Cross-references: UNIPARC:UPI00000D3C05; GB:AE001825; NID:96460670; PIDN:
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0319
A:Map position: 2
C:Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease 12K c

Query Match 43.3%; Score 511; DB 2; Length 228;
Best Local Similarity 45.7%; Pred. No. 1e-11;
Matches 106; Conservative 46; Mismatches 64; Indels 16; Gaps 6;
QY 1 VKLTPKEQKELLYAGVAVRKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
DB 1 MQLTERDRKLLIFTAALARRARRAGLKLHPEAVALLITAEVLEIGIRDGR-----VED 55
QY 61 CMHF-----LKKDEVMPGVGNMVPDLGVATPDGPKLVTVNMPIDPEHFK-AGS--VKF 113
DB 56 LMSFGAAILTPDDVLDGVPELHIEIQVEGTDPDGTCLTVVHDPIFGAASRRVAGBYLLEG 115
QY 114 GCDKDIENLAKGVTELEVTEVNEGPKSLHVGSHFPEAKLKDREKAYGKRLDIPSGN 173
DB 116 G---EIELNAGRPVTLTAVNTADRPVQVSHFPEVFNAGLRFDRAAAYGFRLLNPAGT 172
QY 174 TLIRGAGTRKVLQIPLGGSKVYGMGLVNNIAD-ERHKHAKLDKAKSHGF 224
DB 173 AVRFEPGEEREVDLVLPGSSRTYGMNALVNGDLDAPGTREAALEGARAAAGF 224

RESULT 5
URJB
urease (EC 3.5.1.15) - jack bean
N:Alternate names: urea amidohydrolase
C:Species: Canavalia ensiformis (jack bean)
C:Date: 28-Dec-1987 #sequence_revision 30-Sep-1993 #text_change 05-Oct-2004
C:Accession: JCL1396; S01021; J70001; A60883; S03364; J50642
R:Riddles, P.W.; Whan, V.; Blakeley, R.L.; Zerner, B.
Gene 121, 399-400, 1992
A:Reference number: JCL1396; MUID:93077062; PMID:1446838
A:Contents: erratum
A:Accession: JCL1396
A:Molecule type: mRNA
A:Residues: 1-840 <RID>
A:Cross-references: UNIPROT:P07374; UNIPARC:UPI0000137DA7; GB:I65260
R:Riddles, P.W.; Whan, V.; Blakeley, R.L.; Zerner, B.
Gene 108, 265-267, 1991
A:Title: Cloning and sequencing of a jack bean urease-encoding cDNA.
A:Reference number: J50642; MUID:92084145; PMID:1721034
A:Contents: annotation
A:Note: this sequence has been revised in reference JCL1396
R:Takishima, K.; Suga, T.; Mamiya, G.
Eur. J. Biochem. 175, 151-165, 1988
A:Title: The structure of jack bean urease. The complete amino acid sequence, limited pr
A:Reference number: S01021; MUID:98296463; PMID:3402446
A:Accession: S01021
A:Molecule type: protein
A:Residues: 1-246 'R', 248-840 <TAI>
A:Cross-references: UNIPARC:UPI0000172D9E
R:Mamiya, G.; Takishima, K.; Masakuni, M.; Kayumi, T.; Ogawa, K.; Sekita, T.
Proc. Jpn. Acad. B Phys. Biol. Sci. 61, 395-398, 1985
A:Title: Complete amino acid sequence of jack bean urease.

A:Reference number: J70001
A:Accession: J70001
A:Molecule type: protein
A:Residues: 1-246 'R', 248-257 'P', 259-269 'S', 270-840 <TA2>
A:Cross-references: UNIPARC:UPI0000172D9F
R:Mamiya, G.; Takishima, K.; Masakuni, M.; Kayumi, T.; Ogawa, K.
J. Molec. Biol. 198, 55-59, 1987
A:Title: Complete amino acid sequence of jack bean urease.
A:Reference number: A60883
A:Accession: A60883
A:Molecule type: protein
A:Residues: 1-246 'R', 248-840 <TA3>
A:Cross-references: UNIPARC:UPI0000172D9E
R:Takishima, K.; Mamiya, G.
Protein Seq. Data Anal. 1, 103-106, 1987
A:Title: Location of the essential cysteine residue of jack bean urease.
A:Reference number: S03364; MUID:88190054; PMID:3447159
A:Accession: S03364
A:Molecule type: protein
A:Residues: 591-637 <TA4>
A:Cross-references: UNIPARC:UPI0000172DA0
C:Comment: Each chain of the hexamer binds two nickel ions.
C:Superfamily: urease; urease 11K chain homology; urease 12K chain homology; urease 62K c
C:Keywords: homohexamer; hydrolase; metalloprotein; nickel
F:1-101/Domain: urease 11K chain homology <U11>
F:134-230/Domain: urease 12K chain homology <U12>
F:274-823/Domain: urease 62K chain homology <U62>
F:407,409,490,633/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted
F:490,519,545/Binding site: nickel 1 (Lys, His, His) #status predicted
F:490/Binding site: carbon dioxide (Lys) (covalent) (by urease activase) #status predicted

Query Match 39.1%; Score 462; DB 1; Length 840;
Best Local Similarity 41.4%; Pred. No. 1.5e-08;
Matches 110; Conservative 45; Mismatches 68; Indels 43; Gaps 13;
QY 1 VKLTPKEQKELLYAGVAVRKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
DB 1 MKLSPREVEKGLHNAAGYLAQRLARGVRLNTEAVALLTASQIMYARDGKTKVAQLM-- 58
QY 61 CM--HFLKKDEVMPGVGNMVPDLGVATPDGPKLVTVNMPI--EPDEHFK--GEV--- 111
DB 59 CLGQHLLGRQVLPAVPHLLNAVQVEATPDGPKLVTVHDPIRENGELQALFGSLPLV 118
QY 112 ----KF-----G---C-DKDIENLAKGVTELEVTEVNEGPKSLHVGSHFPEAK 153
DB 119 PSLDKFAETKEDNRIPGTEILCECLTLNIGRAVILKVTSGDRPIQVGVSHYFIEVNP 178
QY 154 ALKFDREKAYGKRLDIPSGNTLRIGAGQTRKVLQIPLGGSKKVIQMN----GLVN--NIA 207
DB 179 YLTFDRRKAYGMRLNIAAGTAVRFPFGDCSKSVTLVSEGNKVIKRGNAIADGPNVETNLE 238
QY 208 DERH----K---HKALDKAKSHGF 226
DB 239 AMHVAVRKSGFGHEE-EKDASEGFTK 263

RESULT 6
A96699
probable urease F12B7.10 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 05-Oct-2004
C:Accession: A96699
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Deary, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: A96699
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-838 <STO>
A;Cross-references: UNIPARC:UPI00000ABA6F; GB:AE005173; NID:G6382524; PIDN:AAF07806.1; C
C;Genetics:
A;Gene: F12B7.10
A;Map position: 1
C;Superfamily: urease; urease 11K chain homology; urease 12K chain homology; urease 62K

Query Match 36.2%; Score 427; DB 2; Length 838;
Best Local Similarity 38.0%; Pred. No. 3e-07;
Matches 100; Conservative 47; Mismatches 73; Indels 43; Gaps 9;
Qy 1 VKLTPKQEKFLYYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLLPRIEXKLELHQAGFLAQKRLARGIRLNYTEAVALLAQILEFIRIDGKSVLAELMDI 60
Qy 61 CMHFLKDEVMVPGVGNMVPDLGVEATPDGKLVTVNWPI-----E----- 101
Db 61 GROLLGRQVLPVALLHLYTVQVEGTFRDGKLVTVHEPISLENGNLELALHGSFLPVPS 120
Qy 102 -----PDH-----FKAGEVKFGDKDIENAGKEVTELEVNEGPKSLHVGSHFHFPEANKA 154
Db 121 LDKPPEVHEGVIIIPGDMKYG-DGSIINHGKRAVVLKVNTGDRPVQVGSYHFIEVNPL 179
Qy 155 LKFDREKAYGKRLDIPSGNTLIRIGAGOTRKVQLIPLGSGKKVIGMGLVNNIAD----- 208
Db 180 LVFDRKALGNLNIAGTAVRFPFEGKSVVIVNIGNKVIRGNGIIVGLVDVDDVNTV 239
Qy 209 -----ERH--KH-KALDKAKSHG 223
Db 240 LMETMERRGFKHLEDID-A-SEG 260

RESULT 7

T29055
urease (EC 3.5.1.5) beta/gamma chain - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 06-Jan-2003
C;Accession: T29055
R;Redenbach, M.; Kieser, H.M.; Denapate, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopw
Mol. Microbiol. 21, 77-96, 1996
A;Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S
A;Reference number: 220556; MUID:97000351; PMID:8843436
A;Accession: T29055
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-227 <RED>
A;Cross-references: UNIPARC:UPI00000DADF4; EMBL:AL031124; PIDN:CAA19973.1
C;Genetics:
A;Gene: ureAB
C;Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease 12K c
C;Keywords: hydrolase

Query Match 35.2%; Score 416; DB 2; Length 227;
Best Local Similarity 40.5%; Pred. No. 3.7e-08;
Matches 94; Conservative 46; Mismatches 76; Indels 16; Gaps 8;
Qy 1 VKLTPKQEKFLYYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MRLTPTERDRLLPGAELARARRARGRLNVPEATALIADTVCEAARDGAR-LAQATER 59
Qy 61 CMHFLKDEVMVPGVGNMVPDLGVEATPDGKLVTVNWPIE-----PDHFKAGEVKFGCD 116
Db 60 ARSVLGPDDVLPVADVTVVHVAVFDDGSRSLAVVADPVGGGLGDD--APGALLPHGD 117
Qy 117 KDIELNAGKEVTELEVNEG--PKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNT 174
Db 118 RP-EPEA---ALRLPVNTATVPVS--VTSFHFPEANPRDLDGRAYGNRLAVPAGSS 171
Qy 175 LRIGAGOTRKVQLIPLGSGKKVIGMGLVNNIAD--ERHKHAKLDKAKSHGFI 225
Db 175 LRIGAGOTRKVQLIPLGSGKKVIGMGLVNNIAD--ERHKHAKLDKAKSHGFI 225

Db 172 VRFGGERVEVGLVPIGGARVAIGFAGLVDPAGAREEARLRAACGYL 223

RESULT 8

T37939
urease (EC 3.5.1.5) - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 12-Jul-2004
C;Accession: T37939
R;McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Pohl, T.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21755
A;Accession: T37939
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-835 <MCD>
A;Cross-references: UNIPROT:O00084; UNIPARC:UPI0000137DA9; EMBL:AL109820; PIDN:CAB52575.1
A;Experimental source: strain 972h-; cosmid c1952
C;Genetics:
A;Gene: SPDB:SPAC1952.11c
A;Map position: 1
C;Superfamily: urease; urease 11K chain homology; urease 12K chain homology; urease 62K c
C;Keywords: hydrolase

Query Match 34.9%; Score 412; DB 2; Length 835;
Best Local Similarity 37.4%; Pred. No. 1.1e-06;
Matches 95; Conservative 39; Mismatches 87; Indels 33; Gaps 7;
Qy 3 LTPKQEKFLYYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEECM 62
Db 1 MQPRELHLTLHLQGLSLAQKRLCRGKLNKLEATSLIASIQIYVYRDGNHNSVADLSLKG 60
Qy 63 HFLKXDEVMPGVGNMVPDLGVEATPDGKLVTVNWPI-EPD---EH----- 105
Db 61 DMLGKHVQPNVLLHLEIMIEATPDGTYLTITDHPICCTDGNLEHALYGSFLPTPSQE 120
Qy 106 -FKAGEVKF-----G-----CDKIEL--NAGKEVTELEVNEGPKSLHVGSHFHFPEA 151
Db 121 LPLPEEKLYAPENSPGFVELEGTELLPNLPR--TPIEVNMGDRPIQVGSYHFIEF 178
Qy 152 NKALFDEKAYGKRLDIPSGNTLIRIGAGOTRKVQLIPLGSGKKVIGMGLVNNIADDERH 211
Db 179 NEKLCFDRSKAYGKRLDIPSGTAIRFEPGVMKIVNLIPIGGAKLIQGGNSLKGVFDDSR 238
Qy 212 KHKALDKAKSHGFI 225
Db 239 TREIVDNLKQGF 252

RESULT 9

A36138
urease (EC 3.5.1.5) 11K chain [validated] - Klebsiella pneumoniae
N;Alternate names: urea amidohydrolase; urease chain A; urease gamma chain
C;Species: Klebsiella pneumoniae
C;Date: 30-Nov-1990 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C;Accession: A36138; S32938
R;Mulrooney, S.B.; Hausinger, R.P.
J. Bacteriol. 172, 5837-5843, 1990
A;Title: Sequence of the Klebsiella aerogenes urease genes and evidence for accessory pr
A;Reference number: A36138; MUID:91008957; PMID:2211515
A;Accession: A36138
A;Molecule type: DNA
A;Residues: 1-100 <MUL>
A;Cross-references: UNIPROT:P18316; UNIPARC:UPI00001103C0; GB:M36068; NID:g149335; PIDN:?
R;Collins, C.M.; Gutman, D.M.; Laman, H.
Mol. Microbiol. 8, 187-198, 1993
A;Title: Identification of a nitrogen-regulated promoter controlling expression of Klebsi
A;Reference number: S32937; MUID:93268094; PMID:8497192
A;Accession: S32938
A;Molecule type: DNA
A;Residues: 1-66, 'K', 68-100 <COL>
A;Cross-references: UNIPARC:UPI0000137D9A; EMBL:L07039; NID:g149330; PIDN:AAA25147.1; PII
R;Jabri, E.; Karplus, P.A.


```

Db      1 MKLTPVEQEKLLIPAAAGELAKQKARGVLLNYPAAAYITCFIMEGARGD-KGVAELMEA 59
QY      61 CMHFL-KKDEVMFGVGNMVPDLGVEATFPDGTGLVTVMNPIDPDEHFKAGEVK 112
        :|||:|||: : ||||| ||||| :||| :|||
Db      60 GRHVLTEKD-VMGVPPEMLDSIQVEATFPDGTGKLVTVHQPI-----SAEVK 104
        :|||:|||: : ||||| ||||| :||| :|||

RESULT 15
A36950
urease (EC 3.5.1.5) 11k chain - Bacillus sp. (strain TB-90)
N:Alternate names: ureA protein
C:Species: Bacillus sp.
C:Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004
C:Accession: A36950
R:R:Maeda, M.; Hidaka, M.; Nakamura, A.; Masaki, H.; Uozumi, T.
J. Bacteriol. 176, 432-442, 1994
A:Title: Cloning, sequencing, and expression of thermophilic Bacillus sp. strain TB-90 u
A:Reference number: A36950; MUID:94117379; PMID:8288539
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <NAE>
A:Cross-references: UNIPROT:Q07399; UNIPARC:UPI0000137D95; GB:D14439; NID:g393296; PIDN:I:
C:Superfamily: urease, gamma subunit; urease 11k chain homology
C:Keywords: hydrolase
F:1-100/Domain: urease 11k chain homology <U11>

Query Match      23.4%; Score 276; DB 2; Length 100;
Best Local Similarity 55.0%; Pred. No. 0.00095;
Matches 55; Conservative 19; Mismatches 25; Indels 1; Gaps 1;

QY      1 VKLTPKEQEKFLYYAGEVAKRKKAEGKLKLNQPEAIAIYISAHIMDEARRGKKTVAQLMEE 60
        :|||:|||: : ||||| ||||| :||| :|||
Db      1 MKLTSREMEKLMIVVAADLARRRKERGLKLNYPEAVAMITYEVLEGARDG-KTVAQLMQY 59
        :|||:|||: : ||||| ||||| :||| :|||

QY      61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGLVTVMNPFI 100
        |:|||:|||: : ||||| ||||| :||| :|||
Db      60 GATILTKEDEVMEGVAEMIPDIQIEATFPDGTGKLVTVHDPFI 99
        |:|||:|||: : ||||| ||||| :||| :|||

Search completed: August 9, 2006, 21:29:40
Job time : 41 secs

```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:20:41 ; Search time 300 Seconds

(without alignments)
696.845 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLYYAGEVA.....ADERHKHAKDKAKSHGFIK 226

Scoring table: BLOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	665	56.3	234	1 UREA_HELBI	Q8q988 helicobacte
2	638	54.0	234	1 UREA_HELHE	P42822 helicobacte
3	619	52.4	238	1 UREA_HELHP	Q9xm24 helicobacte
4	617	52.2	238	2 O8RNU5_HELHP	O8Rnu5 helicobacte
5	617	52.2	238	2 Q93NJ1_HELHP	Q93nj1 helicobacte
6	616	52.2	238	1 UREA_HELPI	P14916 helicobacte
7	616	52.2	238	2 Q93NI3_HELHP	Q93ni3 helicobacte
8	616	52.2	238	2 Q93NI5_HELHP	Q93ni5 helicobacte
9	615	52.1	238	2 Q93NI8_HELHP	Q93ni8 helicobacte
10	614	52.0	238	2 Q93M9_HELHP	Q93m9 helicobacte
11	613	51.9	238	2 Q932S5_HELHP	Q932s5 helicobacte
12	613	51.9	238	2 Q93NI4_HELHP	Q93ni4 helicobacte
13	613	51.9	238	2 Q93J0_HELHP	Q93j0 helicobacte
14	613	51.9	238	2 Q93NJ2_HELHP	Q93nj2 helicobacte
15	611	51.7	238	2 O84F76_HELHP	O84f76 helicobacte
16	610	51.7	238	2 Q93NI9_HELHP	Q93ni9 helicobacte
17	609	51.6	238	2 Q93NH7_HELHP	Q93nh7 helicobacte
18	609	51.6	238	2 Q93NI2_HELHP	Q93ni2 helicobacte
19	608	51.5	238	2 Q93P5_HELHP	Q93p5 helicobacte
20	607	51.4	238	2 Q93NI0_HELHP	Q93ni0 helicobacte
21	607	51.4	238	2 Q93NI1_HELHP	Q93ni1 helicobacte
22	605	51.2	238	2 Q93NH8_HELHP	Q93nh8 helicobacte
23	605	51.2	238	2 Q93NI6_HELHP	Q93ni6 helicobacte
24	604	51.1	238	2 Q93NI7_HELHP	Q93ni7 helicobacte
25	602	51.0	238	2 Q93NH9_HELHP	Q93nh9 helicobacte
26	572	48.4	237	1 UREA_HELFE	O8q715 helicobacte
27	561	47.5	347	2 O7PED9_ANOGA	O7ped9 anopheles g
28	556	47.1	225	1 UREA_HELHP	Q93pj5 helicobacte
29	554	46.9	213	1 UREA_HELHU	P50044 helicobacte
30	529	44.8	219	1 URE23_SULTO	Q972v9 sulfolobus
31	511	43.3	228	1 URE23_DBIRA	Q9ryj3 deinococcus

32	505	42.8	231	1 URE23_PSESM	Q883f3 pseudomonas
33	503	42.6	223	2 O5FB24_CAMLA	O5fb24 campylobact
34	502	42.5	231	2 Q4ZUD1_PSEU2	Q4zud1 pseudomonas
35	484	41.0	231	2 Q35P06_9BRAD	Q35p06 bradyrhizob
36	467	39.5	838	2 Q7XAC5_SOVEN	Q7xac5 glycine max
37	462	39.1	840	1 UREA_CANEN	P07374 canavalia e
38	449	38.0	176	2 O8KIY9_HELHE	O8kiy9 helicobacte
39	448	37.9	177	2 O8KJ05_HELHE	O8kj05 helicobacte
40	446	37.8	172	2 O8KHJ0_HELHE	O8khj0 helicobacte
41	446	37.8	179	2 O8KI21_HELHE	O8ki21 helicobacte
42	445	37.7	185	2 Q9R8R4_HELHP	Q9r8r4 helicobacte
43	443	37.5	185	2 Q9R8R2_HELHP	Q9r8r2 helicobacte
44	439	37.2	164	2 O8KIZ9_HELBI	O8kiz9 helicobacte
45	437	37.0	190	2 Q61616_CAMLA	Q61616 campylobact

ALIGNMENTS

RESULT 1
UREA_HELBI STANDARD; PRT; 234 AA.
AC Q8GH98;
DT 16-JAN-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-MAR-2003, sequence version 1.
DT 21-FEB-2006, entry version 20.
DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase alpha subunit).
GN Name=urea;
OS Helicobacter bizzozeronii.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OC NCBI_TaxID=56877;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=22540228; PubMed=12652903; DOI=10.1080/1042517021000039230;
RA Zhu J., Teng C.H., Chang C.P., Chang C.D., Simpson K.W., Wei C.,
RA McDonough P., McDonough S., Chang Y.F.;
RT "Cloning and characterization of a Helicobacter bizzozeronii urease
gene cluster.";
RL DNA Seq. 13:321-331(2002).
CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -!- SUBUNIT: Heterodimer of an alpha subunit and a beta subunit (By
similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: In the N-terminal section; belongs to the urease gamma
subunit family.
CC -!- SIMILARITY: In the C-terminal section; belongs to the urease beta
subunit family.
CC -!- CAUTION: In Helicobacter the alpha subunit is what is known, in
other bacteria as the beta subunit.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; AF330621; AA015373.1; -; Genomic DNA.
HSSP; P14916; 1E9Z.
SMR; Q8GH98; 1-224.
HAMAP; MF_00739; fused; 1.
InterPro; IPR002019; Urease_beta.
InterPro; IPR002026; Urease_gamma_reg.
InterPro; IPR008223; Urease_gamma_beta.
Pfam; PF00699; Urease_beta; 1.
Pfam; PF00547; Urease_gamma; 1.
PIRSF; PIRSF001225; Urease_gamma_beta; 1.
ProDom; PD002326; Urease_beta; 1.
ProDom; PD002319; Urease_gamma; 1.
TIGRPFAMs; TIGR00192; urease_beta; 1.
TIGRPFAMs; TIGR00193; urease_gamma; 1.
Hydrolase.
CHAIN 1 234 Urease alpha subunit.
REGION 1 102 /FTID=PRO_0000098069.
REGION 103 234 Urease gamma.
REGION 103 234 Urease beta.

```

/FTID=PRO_0000098071.
FT REGION 1 102
FT REGION 103 234
FT REGION 103 234
SQ SEQUENCE 234 AA; 25746 MW; CD5889E3D789C7F8 CRC64;

Query Match 54.0%; Score 638; DB 1; Length 234;
Best Local Similarity 54.8%; Pred. No. 2.2e-16;
Matches 126; Conservative 41; Mismatches 55; Indels 8; Gaps 5;

QY 1 VKLTPEQSEKFLLYAGGEVARKKASGLKLNQPEAJAYISAHIMDEARRGKKTVAOLMEE 60
DB 1 MKLTTPKELDKMLHYAGELAKQKAKGKILNYTEVALISAHVWEARAKKSVADLMQE 60
QY 61 CMHFLKDEVMPCGVGNVPLDGVGEATFPDGTKLVTNVNPIEP--DEHFKAGEV--KFGCDK 117
DB 61 GRTLLKADDVMPGVAHMEVGEIAGFPDGTGLVTIHTPVEAGSDKLAPGEVILK---NE 117
QY 118 DIELNAGKEVTELEVTNPGKSLHVGSHHPFPFANKALKFDREKAYGKRLDIPSGNTLRI 177
DB 118 DIETLNAGKHAVALQKNGKDRPVQVSGSHHPFPFVFNKLLDFDREKAYGKRLDIASGTAVR 177
QY 178 GAGOTRKVOLLPLGGSKKVTGMNGLVNNIADERHKHKALDKAK-SH-GFI 225
DB 178 EPGEEKTVELDIGGNKRYGFNALVDRQADHGKLLAKRAEKHFGTI 227

RESULT 3
UREA HELPU
UREA HELPUJ STANDARD; PRT; 238 AA.
Q9ZM24;
AC 30-MAY-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-MAY-1999, sequence version 1.
DT 07-MAR-2006, entry version 37.
DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase alpha subunit).
GN Name-ureA; Synonyms=hpuA; OrderedLocusNames=JHP0068;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OC NCBI_TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=99210557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -1- SUBUNIT: Heterodimer of an alpha subunit and a beta subunit (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -1- SIMILARITY: In the N-terminal section; belongs to the urease gamma
CC subunit family.
CC -1- SIMILARITY: In the C-terminal section; belongs to the urease beta
CC subunit family.
CC -1- CAUTION: In Helicobacter the alpha subunit is what is known, in
CC other bacteria as the beta subunit.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; AE001446; AD005652.1; -; Genomic_DNA.
DR PIR; B71977; B71977.
DR HSSP; P14916; LE9Z.
DR SWR; Q9ZM24; 1-238.
DR GenomeReviews; AE001439 GR; JHP0068.
DR BioCyc; HPY185963.JHP0068-MONOMER; -.
DR InterPro; IPR002019; Urease_beta.
DR HAMAP; MF_00739; fused; 1.
DR InterPro; IPR002026; Urease_gamma_reg.
DR

```

	Query Match	52.2%	Score 617;	DB 2;	Length 238;
	Best Local Similarity	52.7%	Pred. No. 1.5e-15;		
	Matches 118; Conservative	43;	Mismatches 62;	Indels 1;	Gaps 1
Qy	1 VKLTPQEKEFLYYAGEVARKKAEGKLINOPEAIAIYISAHIMEAREGKKTTVAOLMEE	60			
Dd	1 MKLTFPKSLDKMLHYAGELAKKRKEKGHKLNVEAVRLISAHIMEAREGKKTAABLMOE	60			
Qy	61 CMHFLLKXDEVNPGVGNMVPDLGVZATFPDGTGLVTVNWPIEPDEHPKAGEVFGCDKDIE	120			
Dd	61 GRTLLKPDVDMDGVASMIHEVGIEAMFDDGTGLVTVHTPIANGKLVLPSEL-FLKNEDIT	119			
Qy	121 LNACGEVTELEVNTNEGPKSLHVGSHPHFFPEANKALKPDREKAYGRLDITPSNGNTLRIGAG	180			
Dd	120 INEGKAVSVKVKVGDGPVOIGSHFEHFEVNRCLDFDRKFTFKGLDIASGTAVRFPFG	179			

```

RESULT 5
Q93NJ1_HELPY
ID Q93NJ1_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NJ1;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Urease protein Urea.
GN Name=urea;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RL NUCLEOTIDE SEQUENCE.
RP STRAIN=93K;
RC MEDLINE=23001021; PubMed=14638992;
RX Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RL Taiwan strains";
RL Microbiol. Immunol. 47:813-821 (2003).
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
EMBL: AF373558; AAK69724.1; -; Genomic_DNA.
DR HSSP: P14916; IE9Y.
DR SMR; Q93NJ1; 1-238.
DR GO: 0005737; C: Cytoplasm; IEA.
DR GO: 0016151; F: Nickel ion binding; IEA.
DR GO: 0009039; F: Urease activity; IEA.
DR GO: 0006807; P: Nitrogen compound metabolism; IEA.
DR InterPro: IPR002019; Urease beta.

```

Smith H.O., Fraser C.M., Venter J.C.;
 "The complete genome sequence of the gastric pathogen *Helicobacter pylori*."
 Nature 388:539-547(1997).
 [4]
 NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RC STRAIN=HPK5;
 RC Akada J.K., Shirai M., Takeuchi H., Teuda M., Nakazawa T.;
 "The urase operon in *Helicobacter pylori* is regulated by decay of mRNA."
 Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 [5]
 PROTEIN SEQUENCE OF 1-20.
 MEDLINE=90264448; PubMed=2188975;
 Dunn B.E., Campbell G.P., Perez-Perez G.I., Blaser M.J.;
 "Purification and characterization of urease from *Helicobacter pylori*."
 J. Biol. Chem. 265:9464-9469(1990).
 [6]
 PROTEIN SEQUENCE OF 1-20.
 MEDLINE=90202165; PubMed=2318539;
 Hu L.T., Mobley H.L.T.;
 "Purification and N-terminal analysis of urease from *Helicobacter pylori*."
 Infect. Immun. 58:992-998(1990).
 [7]
 X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS), AND SUBUNIT STRUCTURE.
 MEDLINE=1267477; PubMed=11373617; DOI=10.1038/88563;
 Ha N.-C., Oh S.-T., Sung J.Y., Cha K.A., Lee M.H., Oh B.-H.;
 "Supramolecular assembly and acid resistance of *Helicobacter pylori* urease."
 Nat. Struct. Biol. 8:505-509(2001).
 -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
 -!- SUBUNIT: Heterohexamer of 3 alpha and 3 beta subunits. Four heterohexamers assemble to form a 16 nm spherical complex.
 -!- SUBCELLULAR LOCATION: Cytoplasm.
 -!- MISCELLANEOUS: Oligomerization may protect the enzyme against denaturation in an acidic environment.
 -!- SIMILARITY: In the N-terminal section; belongs to the urease gamma subunit family.
 -!- SIMILARITY: In the C-terminal section; belongs to the urease beta subunit family.
 -!- CAUTION: In *Helicobacter* the alpha subunit is what is known, in other bacteria as the beta subunit.
 Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 Distributed under the Creative Commons Attribution-NoDerivs License
 EMBL: X17079; CAA34932.1; -; Genomic DNA.
 EMBL: M60398; AAA25020.1; -; Genomic DNA.
 EMBL: A0816; CAA0810.1; -; Unassigned DNA.
 EMBL: A07396; CAA00662.1; -; Unassigned DNA.
 EMBL: A800511; AAD07144.1; -; Genomic DNA.
 EMBL: A8032429; BAA84532.1; -; Genomic DNA.
 PIR: A38537; URKCAP.
 PDB: 1E9Y; X-ray; A=1-238.
 PDB: 1E9Z; X-ray; A=1-238.
 GenomesReviews; AE000511_GR; HP0073.
 TIGR: HP0073; -;
 HAMAP; MF_00739; fused; 1.
 InterPro; IPR002019; Urease_beta.
 InterPro; IPR002026; Urease_gamma_reg.
 InterPro; IPR008223; Urease_gamma_beta.
 Pfam; PF00699; Urease_beta; 1.
 Pfam; PF00547; Urease_gamma; 1.
 PIRSF; PIRSF01225; Urease_gamma_beta; 1.
 ProDom; PD002326; Urease_beta; 1.
 ProDom; PD002319; Urease_gamma; 1.
 TIGRFAMs; TIGR00192; urease_beta; 1.
 TIGRFAMs; TIGR00193; urease_gamma; 1.
 3D-structure; Complete proteome; Direct protein sequencing; Hydrolase.
 Urease alpha subunit.
 /FTId=PRO_0000098075.
 CHAIN 1 238
 FT

```
FT REGION 1 102 Urease gamma.
FT REGION 103 238 Urease beta.
FT CONFLICT 14 H -> S (in Ref. 6).
FT CONFLICT 37 A -> R (in Ref. 2).
FT CONFLICT 49 A -> R (in Ref. 2).
FT CONFLICT 132 KN -> PP (in Ref. 2).
FT STRAND 3 3
FT HELIX 5 25
FT TURN 26 26
FT STRAND 30 30
FT STRAND 32 30
FT HELIX 32 49
FT TURN 50 50
FT STRAND 52 52
FT HELIX 54 63
FT STRAND 64 64
FT TURN 67 69
FT STRAND 70 70
FT TURN 72 73
FT HELIX 74 77
FT STRAND 80 87
FT TURN 88 89
FT STRAND 90 97
FT STRAND 100 100
FT STRAND 104 105
FT TURN 108 109
FT STRAND 111 112
FT STRAND 115 115
FT STRAND 117 119
FT TURN 120 123
FT STRAND 128 133
FT STRAND 135 137
FT STRAND 139 142
FT TURN 143 144
FT STRAND 145 145
FT STRAND 147 149
FT HELIX 147 149
FT TURN 150 150
FT TURN 152 153
FT STRAND 154 156
FT STRAND 158 161
FT TURN 162 163
FT STRAND 164 166
FT STRAND 168 169
FT TURN 170 171
FT STRAND 173 176
FT TURN 178 179
FT STRAND 181 188
FT TURN 191 192
FT STRAND 194 195
FT STRAND 197 197
FT TURN 198 199
FT STRAND 200 201
FT STRAND 204 205
FT STRAND 207 207
FT STRAND 208 221
FT TURN 222 222
FT TURN 224 225
FT STRAND 230 230
FT STRAND 236 236
SQ SEQUENCE 238 AA; 26540 MW; 4E77328669CD9A2D CRC64;

Query Match 52.2%; Score 616; DB 1; Length 238;
Best Local Similarity 52.7%; Pred. No. 1.6e-15;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

Qy 1 VKLTPKQEKFLYYAGEVARKKAEGLKNQPEAIYISAHIMDEARAGKKTVAOLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRKEGIKLNYVEAVALISAHIMEARAGKKTAAELMQE 60

Qy 61 CMHFLKKDEVMGPNMVPDLGVEATFPDGTGLVTVMNPIEPDHFHFKAGEVKGCDKID 120
Db 61 GRTLKPPDDVMDGVASMIHEVGIEAMFPDGTGLVTVHTPIEANGKLVPGEL-FLKNEDIT 119

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180

Qy 120 INEGKAVSVKXNVGDRPVQIGSHFHFPEVNRVYDFDREKTFGKELDIASGTAVRPEPG 179
Db 120 INEGKAVSVKXNVGDRPVQIGSHFHFPEVNRVYDFDREKTFGKELDIASGTAVRPEPG 179

Qy 191 QTRKVLQPLGSKVKVIGMGNLVNNIADERHKHAKLDKAKSHGF 224
Db 191 QTRKVLQPLGSKVKVIGMGNLVNNIADERHKHAKLDKAKSHGF 224

Qy 180 BEKSVELIDIGGSRRIFFGNALVDRQADNESKCIALHRAKERGF 223
Db 180 BEKSVELIDIGGSRRIFFGNALVDRQADNESKCIALHRAKERGF 223

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180
```

Db 120 INEGKAVSVKXNVGDRPVQIGSHFHFPEVNRVYDFDREKTFGKELDIASGTAVRPEPG 179

Qy 181 QTRKVLQPLGSKVKVIGMGNLVNNIADERHKHAKLDKAKSHGF 224

Db 180 BEKSVELIDIGGSRRIFFGNALVDRQADNESKCIALHRAKERGF 223

RESULT 7

Q93NI3 HELPY PRELIMINARY; PRT; 238 AA.

AC Q93NI3;

DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.

DT 01-DEC-2001, sequence version 1.

DT 07-FEB-2006, entry version 20.

DE Urease protein UreA.

GN Name=ureA;

OS Helicobacter pylori (Campylobacter pylori).

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

OC Helicobacteraceae; Helicobacter.

OX NCBI_TaxID=210;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=52;

RX MEDLINE=23001021; PubMed=14638992;

RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;

RT "Rate of Helicobacter pylori infection in children and clonality of

RL Taiwan strains.";

RL Microbiol. Immunol. 47:813-821(2003).

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

CC EMBL; AF373570; AAK69736.1; -; Genomic_DNA.

DR HSP; P14916; 1E9Y.

DR SMR; Q93NI3; 1-238.

DR GO; GO:0005737; Cytoplasm; IEA.

DR GO; GO:0016151; F:nickel ion binding; IEA.

DR GO; GO:0009039; F:urease activity; IEA.

DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.

DR InterPro; IPR002019; Urease_beta.

DR InterPro; IPR002026; Urease_gamma_reg.

DR InterPro; IPR008223; Urease_gamma_beta.

DR Pfam; PF00699; Urease_beta; 1.

DR PIRSF; PIRSF001225; Urease_gamma; 1.

DR ProDom; PD002326; Urease_beta; 1.

DR ProDom; PD002319; Urease_gamma; 1.

DR TIGRFAMs; TIGR00192; urease_beta; 1.

DR TIGRFAMs; TIGR00193; urease_gamma; 1.

SQ SEQUENCE 238 AA; 26587 MW; DFD8A17D016FC351 CRC64;

Query Match 52.2%; Score 616; DB 2; Length 238;

Best Local Similarity 53.1%; Pred. No. 1.6e-15;

Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;

Qy 1 VKLTPKQEKFLYYAGEVARKKAEGLKNQPEAIYISAHIMDEARAGKKTVAOLMEE 60

Db 1 MKLTPKELDKMLHYAGELAKRKEGIKLNTVEAVALISAHIMEARAGKKSAAELMQE 60

Qy 61 CMHFLKKDEVMGPNMVPDLGVEATFPDGTGLVTVMNPIEPDHFHFKAGEVKGCDKID 120

Db 61 GRTLKPPDDVMDGVASMIHEVGIEAMFPDGTGLVTVHTPIEANGKLVPGEL-FLKNEDIT 119

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180

Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180

Qy 120 INEGKAVSVKXNVGDRPVQIGSHFHFPEVNRVYDFDREKTFGKELDIASGTAVRPEPG 179

Db 120 INEGKAVSVKXNVGDRPVQIGSHFHFPEVNRVYDFDREKTFGKELDIASGTAVRPEPG 179

Qy 191 QTRKVLQPLGSKVKVIGMGNLVNNIADERHKHAKLDKAKSHGF 224

Db 191 QTRKVLQPLGSKVKVIGMGNLVNNIADERHKHAKLDKAKSHGF 224

Qy 180 BEKSVELIDIGGSRRIFFGNALVDRQADNESKCIALHRAKERGF 223

Db 180 BEKSVELIDIGGSRRIFFGNALVDRQADNESKCIALHRAKERGF 223


```
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RL Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AF373580; AAK69746.1; -; Genomic_DNA.
DR HSSP; P14916; IE9Y.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016511; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR Pfam; PF00699; Urease_beta; 1.
DR ProDom; PD002326; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
DR SEQUENCE 238 AA; 26600 MW; 64A0C32A043A88D8 CRC64;
SQ
Query Match 52.08; Score 614; DB 2; Length 238;
Best Local Similarity 52.78; Pred. No. 2e-15;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;
QY 1 VKLTPKEQKFLLYAGVAVKRAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELARKKEKIGKLNVEAVALISAHIMEARAGKKSAAELMQE 60
QY 61 CMHFLKDEVMGVMGNNVDPDLGVEATPDGTKLVTVNWPIDPEHFKAGEVKFGCDKDIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELVNTEGPKSLHVGSHFFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 INEGKAVSVKVNVDGVPVQIGSHFFPEVNRVYLDREKTFGKRLDIASGTAVRPEPG 179
QY 181 QTRKVLQIPLGGSKKVGNNGLVNNIADERHKHKALDKAKSHGF 224
Db 180 EKSVELIDIGNRRIFGNALVDRQADNESKKIALHRAKERGF 223
RESULT 11
Q932S5_HELPY
ID Q932S5_HELPY PRELIMINARY; PRT; 238 AA.
AC Q932S5;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Urease protein Urea.
GN Name=urea;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=100;
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RL Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RL Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AF373580; AAK69746.1; -; Genomic_DNA.
DR HSSP; P14916; IE9Y.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016511; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR Pfam; PF00699; Urease_beta; 1.
DR ProDom; PD002326; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
DR SEQUENCE 238 AA; 26600 MW; 64A0C32A043A88D8 CRC64;
SQ
Query Match 52.08; Score 614; DB 2; Length 238;
Best Local Similarity 52.78; Pred. No. 2e-15;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;
QY 1 VKLTPKEQKFLLYAGVAVKRAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELARKKEKIGKLNVEAVALISAHIMEARAGKKSAAELMQE 60
QY 61 CMHFLKDEVMGVMGNNVDPDLGVEATPDGTKLVTVNWPIDPEHFKAGEVKFGCDKDIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELVNTEGPKSLHVGSHFFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 INEGKAVSVKVNVDGVPVQIGSHFFPEVNRVYLDREKTFGKRLDIASGTAVRPEPG 179
QY 181 QTRKVLQIPLGGSKKVGNNGLVNNIADERHKHKALDKAKSHGF 224
Db 180 EKSVELIDIGNRRIFGNALVDRQADNESKKIALHRAKERGF 223
RESULT 11
Q932S5_HELPY
ID Q932S5_HELPY PRELIMINARY; PRT; 238 AA.
AC Q932S5;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Urease protein Urea.
GN Name=urea;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=100, 101, 5c, 85K, 86K, and 89K;
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RL Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RL Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AF373584; AAK69750.1; -; Genomic_DNA.
DR EMBL; AF373557; AAK69723.1; -; Genomic_DNA.
DR EMBL; AF373559; AAK69725.1; -; Genomic_DNA.
DR EMBL; AF373581; AAK69747.1; -; Genomic_DNA.
DR EMBL; AF373582; AAK69748.1; -; Genomic_DNA.
DR EMBL; AF373583; AAK69749.1; -; Genomic_DNA.
DR HSSP; P14916; IE9Y.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016511; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR Pfam; PF00699; Urease_beta; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_gamma; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
DR SEQUENCE 238 AA; 26614 MW; E4A2917D016FD8DB CRC64;
SQ
Query Match 51.9%; Score 613; DB 2; Length 238;
Best Local Similarity 52.7%; Pred. No. 2.1e-15;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;
QY 1 VKLTPKEQKFLLYAGVAVKRAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELARKKEKIGKLNVEAVALISAHIMEARAGKKSAAELMQE 60
QY 61 CMHFLKDEVMGVMGNNVDPDLGVEATPDGTKLVTVNWPIDPEHFKAGEVKFGCDKDIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELVNTEGPKSLHVGSHFFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 INEGKAVSVKVNVDGVPVQIGSHFFPEVNRVYLDREKTFGKRLDIASGTAVRPEPG 179
QY 181 QTRKVLQIPLGGSKKVGNNGLVNNIADERHKHKALDKAKSHGF 224
Db 180 EKSVELIDIGNRRIFGNALVDRQADNESKKIALHRAKERGF 223
RESULT 12
Q93NI4_HELPY
ID Q93NI4_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NI4;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Urease protein Urea.
GN Name=urea;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=48;
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RL Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AF373569; AAK69735.1; -; Genomic_DNA.
DR HSSP; P14916; IE9Y.
DR SMR; Q93NI4; 1-238.
DR
```



```
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR InterPro; IPR008223; Urease_gamma_beta.
DR Pfam; PF00699; Urease_beta; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26512 MW; CA6677D2788BDD2E CRC64;

Query Match 51.9%; Score 613; DB 2; Length 238;
Best Local Similarity 52.2%; Pred. No. 2.1e-15;
Matches 117; Conservative 44; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKFLYYAGVARKRAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
DB 1 MKLTPKELDKLMLHYAGELAKRKEKGKLNVEAVALISAHIMEEARAGKKSAAELMQE 60

QY 61 CMHFLKDEVMGPNVMDLGVATFPDGTGLVTNNWPIEDPDEHFKAGVKFGCDKDIE 120
DB 61 GRSLLKPDVMDGVASMIHEVGIEAMPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFRANKALKFDEKAYGKRLDIPSGNLTIRIGAG 180
DB 120 INEGKKAVSVKVNVDGDRPVOIGSHFFHFFVFNRYLDFDREKTFGKRLDIASGTTAVRFPFG 179

QY 181 QTRKVLQIPLOGSKKVGNGLVNNIADRHKHKALDKAKSHGF 224
DB 180 BEKSVELIDIGNRRIFGFNALVDRQADNESKIALHRAKERGF 223

RESULT 13
Q93NJ0_HELPY
ID Q93NJ0_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NJ0;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Urease protein Urea.
GN Name=urea;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=17c;
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RT Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AF373560; AAK69726.1; -; Genomic_DNA.
CC HSSP; P14916; IE9Y.
CC SMR; Q93NJ0; 1-238.
CC GO; GO:0005737; C:cytoplasm; IEA.
CC GO; GO:0016151; F:nickel ion binding; IEA.
CC GO; GO:0009039; F:urease activity; IEA.
CC GO; GO:0006807; P:nitrogen compound metabolism; IEA.
CC InterPro; IPR002019; Urease_beta.
CC InterPro; IPR002026; Urease_gamma_reg.
CC Pfam; PF00699; Urease_beta; 1.
CC ProDom; PD002326; Urease_beta; 1.
CC ProDom; PD002319; Urease_gamma; 1.
CC TIGRFAMs; TIGR00192; urease_beta; 1.
CC TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26512 MW; CA6677D2788BDD2E CRC64;
```

```
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26630 MW; 5C082660AAC88469 CRC64;

Query Match 51.9%; Score 613; DB 2; Length 238;
Best Local Similarity 52.7%; Pred. No. 2.1e-15;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKFLYYAGVARKRAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
DB 1 MNLTPKELDKLMLHYAGELAKRKEKGKLNVEAVALISAHIMEEARAGKKSAAELMQE 60

QY 61 CMHFLKDEVMGPNVMDLGVATFPDGTGLVTNNWPIEDPDEHFKAGVKFGCDKDIE 120
DB 61 GRTLKPDVMDGVASMIHEVGIEAMPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFRANKALKFDEKAYGKRLDIPSGNLTIRIGAG 180
DB 120 INEGKKAVSVKVNVDGDRPVOIGSHFFHFFVFNRYLDFDREKTFGKRLDIASGTTAVRFPFG 179

QY 181 QTRKVLQIPLOGSKKVGNGLVNNIADRHKHKALDKAKSHGF 224
DB 180 BEKSVELIDIGNRRIFGFNALVDRQADNESKIALHRAKERGF 223

RESULT 14
Q93NJ2_HELPY
ID Q93NJ2_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NJ2;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Urease protein Urea.
GN Name=urea;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=96C;
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RT Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AF373556; AAK69722.1; -; Genomic_DNA.
CC HSSP; P14916; IE9Y.
CC SMR; Q93NJ2; 1-238.
CC GO; GO:0005737; C:cytoplasm; IEA.
CC GO; GO:0016151; F:nickel ion binding; IEA.
CC GO; GO:0009039; F:urease activity; IEA.
CC GO; GO:0006807; P:nitrogen compound metabolism; IEA.
CC InterPro; IPR002019; Urease_beta.
CC InterPro; IPR002026; Urease_gamma_reg.
CC InterPro; IPR008223; Urease_gamma_beta.
CC Pfam; PF00699; Urease_beta; 1.
CC Pfam; PF00547; Urease_gamma; 1.
CC ProDom; PD002326; Urease_beta; 1.
CC ProDom; PD002319; Urease_gamma; 1.
CC TIGRFAMs; TIGR00192; urease_beta; 1.
CC TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26630 MW; E4A2917D17B5D8DB CRC64;
```

```
Query Match      51.7%; Score 611; DB 2; Length 238;
Best Local Similarity 52.7%; Pred.No. 2.6e-15;
Matches 118; Conservative 42; Mismatches 63; Indels 1; Gaps 1;
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:34:02 ; Search time 196 Seconds
(without alignments)
1324.995 Million cell updates/sec

Title: US-09-904-994b-3

Perfect score: 2999

Sequence: 1 MKMKQEVNTYGTGDKV.....KLCTSKPTSQVLAQRVTFP 568

Scoring table: BLOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_8.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*
- 10: Geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2999	100.0	568	5	ADJ58239 Urease su
2	2973	99.1	568	5	Adj58251 UreaseXY
3	2960	98.7	568	5	ADJ58242 UreaseXY
4	2953	98.5	568	5	ADJ58248 UreaseXY
5	2591	86.4	496	5	ADJ58245 UreaseXY
6	2326	77.6	569	8	ADQ37851 H. bizzoz
7	2297	76.6	569	2	AAR04580 Part of p
8	2297	76.6	569	2	AAR07194 H. pylori
9	2297	76.6	569	2	AAR98511 H. pylori
10	2297	76.6	569	6	ABU30687 Protein e
11	2297	76.6	569	8	ADU05536 H. pylori
12	2282	76.1	568	8	ADG09177 H. pylori
13	2282	76.1	570	4	ABE55116 Hylcobac
14	2281	76.1	569	2	AAR67378 H. pylori
15	2244	74.8	569	2	AAR74337 Helicobac
16	2244	74.8	569	2	AAR06730 H. felis
17	2243	74.8	749	2	AAR80599 Helicobac
18	2238	74.6	806	2	AAR67371 Urease A
19	2237	74.6	559	2	AAR13550 B subunit
20	2229	74.3	559	8	ADM28644 Helicobac
21	2029	67.7	568	2	AAR36387 Urease ga
22	1927	64.3	568	6	ABU19714 Protein e
23	1921	64.1	566	6	ABU41362 Protein e

24	1913	63.8	568	6	ABU21842	Abu21842 Protein e
25	1902	63.4	568	6	ABU22314	Abu22314 Protein e
26	1843	61.5	572	6	ABU30286	Abu30286 Protein e
27	1838	61.3	566	6	ABU38819	Abu38819 Protein e
28	1834	61.2	566	7	ABO78141	ABO78141 Pseudomon
29	1834	61.2	567	6	ABU32083	Abu32083 Protein e
30	1831	61.1	569	7	ADF05744	Adf05744 Bacterial
31	1831	61.1	779	6	ABU41210	Abu41210 Protein e
32	1830	61.0	625	7	ABO62699	ABO62699 Klebsiell
33	1829	61.0	567	6	ADA34625	Ada34625 Acinetoba
34	1829	61.0	568	2	AAR67379	Aar67379 P. mirabi
35	1826	60.9	566	6	ABU17040	Abu17040 Protein e
36	1826	60.9	567	2	AAW37774	Aaw37774 Klebsiell
37	1826	60.9	567	3	AAW81825	Aaw81825 Klebsiell
38	1818	60.6	568	6	ABU28899	Abu28899 Protein e
39	1818	60.6	568	6	ABU28920	Abu28920 Protein e
40	1818	60.6	568	7	ADC00113	Adc00113 Enterohae
41	1810	60.4	567	6	ABU40328	Abu40328 Protein e
42	1804	60.2	570	2	AAR67380	Aar67380 Jack bean
43	1804	60.2	571	4	AAG81512	Aag81512 S. epider
44	1804	60.2	811	6	ABU42650	Abu42650 Protein e
45	1803	60.1	573	5	ABP40181	Abp40181 Staphyloc

ALIGNMENTS

RESULT 1
ADJ58239
ID ADJ58239 standard; protein; 568 AA.
XX
AC ADJ58239;
XX
DT 06-MAY-2004 (first entry)
XX
DE Urease subunit polypeptide Y.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
XX
N-PSDB; ADJ58237.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
the diagnosis of Helicobacter felis infections and in the preparation of
vaccines.
XX
PT Claim 10; SEQ ID NO 3; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
Y subunit polypeptides and immunogenic fragments. The polypeptides are
used in the manufacture of vaccines against Helicobacter felis infections
and in diagnostic tests to detect antibodies against Helicobacter felis.
XX
CC Helicobacter felis is difficult to grow so it is more convenient to use
the expression products of the genes encoding the urease X and Y subunits
in the manufacture of vaccines. The present sequence represents an urease
XX
Y subunit polypeptide of the invention.
XX
SQ Sequence 568 AA;

Query Match 100.0%; Score 2999; DB 5; Length 568;
 Best Local Similarity 100.0%; Pred. No. 3.3e-52;
 Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEKFGAGKTIREGMQSNP 60
 DB 1 MKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEKFGAGKTIREGMQSNP 60

QY 61 DENTLDLVTNAMIIDYTIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTEAL 120
 DB 61 DENTLDLVTNAMIIDYTIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTEAL 120

QY 121 AGEQMIIITAGIDSHTHFLSPQQPPTALANGVTTMFGGCTGPDVGTNATTITPGKWNLHR 180
 DB 121 AGEQMIIITAGIDSHTHFLSPQQPPTALANGVTTMFGGCTGPDVGTNATTITPGKWNLHR 180

QY 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240
 DB 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240

QY 241 DVQVCIIHTDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTIMAGELNIIIPSSTT 300
 DB 241 DVQVCIIHTDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTIMAGELNIIIPSSTT 300

QY 301 PTPIYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIIRPGSIAAEDVLHDMGVIAMTSS 360
 DB 301 PTPIYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIIRPGSIAAEDVLHDMGVIAMTSS 360

QY 361 DSQAMGRAGEVTPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSE 420
 DB 361 DSQAMGRAGEVTPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSE 420

QY 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTQPYYIREMFGH 480
 DB 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTQPYYIREMFGH 480

QY 481 HGKAKFDTSIITFSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKFNDDTKAKITVDPK 540
 DB 481 HGKAKFDTSIITFSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKFNDDTKAKITVDPK 540

QY 541 TFEVFDGKLTCKTSPQVPLAQRYTFF 568
 DB 541 TFEVFDGKLTCKTSPQVPLAQRYTFF 568

RESULT 2

ADJ58251
 ID ADJ58251 standard; protein; 568 AA.

XX AC ADJ58251;

XX DT 06-MAY-2004 (first entry)

XX DE UreaseXY subunit #8.

XX KW Immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.

XX OS Helicobacter felis.

XX EP1176192-A2.

XX PD 30-JAN-2002.

XX PF 11-JUL-2001; 2001EP-00202666.

XX PR 17-JUL-2000; 2000EP-00202565.

XX PA (ALKU) AKZO NOBEL NV.

XX FI Kusters JG, Cattoli G;

XX DR WPI; 2002-124384/17.

XX DR N-PSDB; ADJ58249.

XX

PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in the diagnosis of Helicobacter felis infections and in the preparation of vaccines.

PT

XX Disclosure; SEQ ID NO 15; 76pp; English.

XX

CC The present invention relates to a novel Helicobacter felis urease X and Y subunit polypeptides and immunogenic fragments. The polypeptides are used in the manufacture of vaccines against Helicobacter felis infections and in diagnostic tests to detect antibodies against Helicobacter felis. Helicobacter felis is difficult to grow so it is more convenient to use the expression products of the genes encoding the urease X and Y subunits in the manufacture of vaccines. The present sequence represents an ureaseXY subunit of the invention.

XX

SQ Sequence 568 AA;

Query Match 99.1%; Score 2973; DB 5; Length 568;

Best Local Similarity 98.9%; Pred. No. 1.1e-51;

Matches 562; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEKFGAGKTIREGMQSNP 60

DB 1 MKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEKFGAGKTIREGMQSNP 60

QY 61 DENTLDLVTNAMIIDYTIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTEAL 120

DB 61 DENTLDLVTNAMIIDYTIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTEAL 120

QY 121 AGEQMIIITAGIDSHTHFLSPQQPPTALANGVTTMFGGCTGPDVGTNATTITPGKWNLHR 180

DB 121 AGEQMIIITAGIDSHTHFLSPQQPPTALANGVTTMFGGCTGPDVGTNATTITPGKWNLHR 180

QY 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240

DB 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240

QY 241 DVQVCIIHTDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTIMAGELNIIIPSSTT 300

DB 241 DVQVCIIHTDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTIMAGELNIIIPSSTT 300

QY 301 PTPIYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIIRPGSIAAEDVLHDMGVIAMTSS 360

DB 301 PTPIYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIIRPGSIAAEDVLHDMGVIAMTSS 360

QY 361 DSQAMGRAGEVTPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSE 420

DB 361 DSQAMGRAGEVTPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSE 420

QY 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTQPYYIREMFGH 480

DB 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTQPYYIREMFGH 480

QY 481 HGKAKFDTSIITFSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKFNDDTKAKITVDPK 540

DB 481 HGKAKFDTSIITFSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKFNDDTKAKITVDPK 540

QY 541 TFEVFDGKLTCKTSPQVPLAQRYTFF 568

DB 541 TFEVFDGKLTCKTSPQVPLAQRYTFF 568

RESULT 3

ADJ58242

ID ADJ58242 standard; protein; 568 AA.

XX AC ADJ58242;

XX DT 06-MAY-2004 (first entry)

XX DE UreaseXY subunit #2.

immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.

Helicobacter felis.

EP1176192-A2.

30-JAN-2002.

11-JUL-2001; 2001EP-00202666.

17-JUL-2000; 2000EP-00202565.

(ALKU) AKZO NOBEL NV.

Kusters JG, Cattoli G;

WPI; 2002-124384/17.

N-PSDB; ADJ58240.

Novel Helicobacter felis urease X and Y subunit polypeptides, useful in the diagnosis of Helicobacter felis infections and in the preparation of vaccines.

Disclosure; SEQ ID NO 6; 76pp; English.

The present invention relates to a novel Helicobacter felis urease X and Y subunit polypeptides and immunogenic fragments. The polypeptides are used in the manufacture of vaccines against Helicobacter felis infections and in diagnostic tests to detect antibodies against Helicobacter felis. Helicobacter felis is difficult to grow so it is more convenient to use the expression products of the genes encoding the urease X and Y subunits in the manufacture of vaccines. The present sequence represents an ureaseXY subunit of the invention.

Sequence 568 AA;

Query Match 98.7%; Score 2960; DB 5; Length 568;

Best Local Similarity 98.6%; Pred. No. 2e-51;

Matches 560; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

1 MMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEELKFGAGKTIREGMQSNP 60

1 MMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEELKFGAGKTIREGMQSNP 60

61 DENTLDLVTNAMIIDYTIYKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120

61 DENTLDLVTNAMIIDYTIYKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120

121 AGEEMIITAGGIDSHTHFLSPQQPPTALANGVTTMFGGGTGPVDTNATTITPGKNLHR 180

121 AGEEMIITAGGIDSHTHFLSPQQPPTALANGVTTMFGGGTGPVDTNATTITPGKNLHR 180

181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240

181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240

241 DVQVCIIHTDVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPVITMAGELNLPSTT 300

241 DVQVCIIHTDVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPVITMAGELNLPSTT 300

301 PTPYTTINTVAEHLDMLTCHLDRKIREDLQFSQSRIRPGSIAAEDVLHDMGVIA 360

301 PTPYTTINTVAEHLDMLTCHLDRKIREDLQFSQSRIRPGSIAAEDVLHDMGVIA 360

361 DSOAMGRAGEVIRPTWQADNKKFPGKLPEDGKNDNFRKRYISKYTINPALTHGVSE 420

361 DSOAMGRAGEVIRPTWQADNKKFPGKLPEDGKNDNFRKRYISKYTINPALTHGVSE 420

421 YIGSVEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYVREMGH 480

421 YIGSVEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYVREMGH 480

481 HGKAKFDTSTFVSKVAYENGVEKGLERQVLPVKNCRNITKKDKFNDKTAKITVD 540

481 HGKAKFDTSTFVSKVAYENGVEKGLERQVLPVKNCRNITKKDKFNDKTAKITVD 540

Db 481 HGKAKFDTSTFVSKVAYENGVEKGLERQVLPVKNCRNITKKDKFNDKTAKITVD 540

QY 541 TFEVFDGKLGKTSKPTSQVPLAQRYTFF 568

Db 541 TFEVFDGKLGKTSKPTSQVPLAQRYTFF 568

RESULT 4

ADJ58248

ID ADJ58248 standard; protein; 568 AA.

AC ADJ58248;

DT 06-MAY-2004 (first entry)

UreaseXY subunit #6.

immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.

Helicobacter felis.

EP1176192-A2.

30-JAN-2002.

11-JUL-2001; 2001EP-00202666.

17-JUL-2000; 2000EP-00202565.

(ALKU) AKZO NOBEL NV.

Kusters JG, Cattoli G;

WPI; 2002-124384/17.

N-PSDB; ADJ58246.

Novel Helicobacter felis urease X and Y subunit polypeptides, useful in the diagnosis of Helicobacter felis infections and in the preparation of vaccines.

Disclosure; SEQ ID NO 12; 76pp; English.

The present invention relates to a novel Helicobacter felis urease X and Y subunit polypeptides and immunogenic fragments. The polypeptides are used in the manufacture of vaccines against Helicobacter felis infections and in diagnostic tests to detect antibodies against Helicobacter felis. Helicobacter felis is difficult to grow so it is more convenient to use the expression products of the genes encoding the urease X and Y subunits in the manufacture of vaccines. The present sequence represents an ureaseXY subunit of the invention.

Sequence 568 AA;

Query Match 98.5%; Score 2953; DB 5; Length 568;

Best Local Similarity 98.2%; Pred. No. 2.8e-51;

Matches 558; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

1 MMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEELKFGAGKTIREGMQSNP 60

1 MMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEELKFGAGKTIREGMQSNP 60

61 DENTLDLVTNAMIIDYTIYKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120

61 DENTLDLVTNAMIIDYTIYKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120

121 AGEEMIITAGGIDSHTHFLSPQQPPTALANGVTTMFGGGTGPVDTNATTITPGKNLHR 180

121 AGEEMIITAGGIDSHTHFLSPQQPPTALANGVTTMFGGGTGPVDTNATTITPGKNLHR 180

181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240

181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240

[illegible]

```

RESULT 5
ADJ58245
ADJ58245 standard; protein; 496 AA.
XX
XX ADJ58245;
AC
XX
XX 06-MAY-2004 (first entry)
XX
XX UreaseXY subunit #4.
DE
XX
XX immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
KW
XX
XX Helicobacter felis.
OS
XX
XX EP1176192-A2.
PN
XX
XX 30-JAN-2002.
PD
XX
XX 11-JUL-2001; 2001EP-00202666.
XX
XX 17-JUL-2000; 2000EP-00202565.
PR
XX
XX (ALKU ) AKZO NOBEL NV.
PA
XX
XX Kusters JG, Cattoli G;
PI
XX
XX WPI; 2002-124384/17.
DR
XX
XX N-PSDB; ADJ58243.
DR
XX
XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX
XX Disclosure; SEQ ID NO 9; 76pp; English.
PS
XX
XX The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
XX Sequence 496 AA;
SQ

```

Best Local Similarity 99.2%; Pred. No. 4.3e-44;		
Matches	491; Conservative	1; Mismatches 2; Indels 1; Gaps 17
QY	1	MMKKQEVYNTYPTKGDVRLGDTDLWAEVEHDYTYGELKFGAGKTIREGMGQSNP 60
DB	1	MMKKQEVYNTYPTKGDVRLGDTDLWAEVEHDYTYGELKFGAGKTIREGMGQSNP 60
QY	61	DENTLDLVITNAMIIDYTYGIKADIGIKNGKIHGIGKAGNKMQDGVSPHMVVGVTGTEAL 120
DB	61	DENTLDLVITNAMIIDYTYGIKADIGIKNGKIHGIGKAGNKMQDGVSPHMVVGVTGTEAL 120
QY	121	AGEGMIITAGGIDSHTHFLSPQOFPPTALANGVTTMFGGTGPVDGTTNATTITPGKWNLHR 180
DB	121	AGEGMIITAGGIDSHTHFLSPQOFPPTALANGVTTMFGGTGPVDGTTNATTITPGKWNLHR 180
QY	181	MLRAAEYSNMVGFGLKGNSSSKKQLVQVEAGAIGFKLHEDWGTTTPSAIDHCLLSVADEY 240
DB	181	MLRAAEYSNMVGFGLKGNSSSKKQLVQVEAGAIGFKLHEDWGTTTPSAIDHCLLSVADEY 240
QY	241	DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSDDVITMAGELNILPSSTT 300
DB	241	DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSDDVITMAGELNILPSSTT 300
QY	301	PTIPYNTINTVAEHLDMLTCHLDKRIREDLQFSQRIIRPGSIAAEDVLDHMGVIAMTSS 360
DB	301	PTIPYNTINTVAEHLDMLTCHLDKRIREDLQFSQRIIRPGSIAAEDVLDHMGVIAMTSS 360
QY	361	DSQAMGRAGEVIPRTWTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSE 420
DB	361	DSQAMGRAGEVIPRTWTADKNKKEFGKLPEDGADNDNFRIKRYISKYTTINPALTHGVSE 420
QY	421	YIGSVBEGKIADLVWNPAPFFGVKPKIVIKGMMVFSEMGDSNASVPTQPVVYREMPGH 480
DB	421	YIGSVBEGKIADLVWNPAPFFGVKPKIVIKGMMVFSEMGDSNASVPTQPVVYREMPGH 480
QY	481	HGKAKEDTSTIF-VS 494
DB	481	HGKAKEDTSTIFRS 495
RESULT 6		
ADQ37851		
ID	ADQ37851	standard; protein; 569 AA.
XX	XX	
XX	ADQ37851;	
DT	07-OCT-2004	(first entry)
XX	XX	
XX	XX	H. bizzozeronii ureB polypeptide.
XX	XX	Urease; urease gene cluster; urease structural gene;
KW	KW	urease accessory gene; ureB; Helicobacter bizzozeronii infection;
KW	KW	antibacterial; enzyme.
XX	XX	
XX	OS	Helicobacter bizzozeronii.
XX	XX	
PN	US2004142343-A1.	
XX	XX	
PD	22-JUL-2004.	
XX	XX	
PF	12-AUG-2003;	2003US-00639273.
XX	XX	
PR	16-AUG-2002;	2002US-0404337P.
XX	XX	
PA	(CHAN/) CHANG Y.	
PA	(SIMP/) SIMPSON K W.	
PA	(ZHUJ/) ZHU J.	
XX	XX	
PI	Chang Y, Simpson KW, Zhu J;	
XX	XX	
DR	WPI; 2004-533502/51.	
DR	N-PSDB; ADQ37850.	
XX	XX	

PT Novel isolated nucleic acid molecule having urease gene cluster, and
 PT conferring on Helicobacter bizzozeronii ability to produce urease, useful
 PT as vaccine for preventing disease in mammals infected by H.bizzozeronii.
 XX
 PS
 PS Claim 24; SEQ ID NO 5; 40pp; English.

XX The invention relates to an isolated nucleic acid molecule conferring on
 CC Helicobacter bizzozeronii an ability to produce urease, where the nucleic
 CC acid molecule is a urease gene cluster comprising at least one urease
 CC structural gene and at least one urease accessory gene. The nucleic acid
 CC molecule is chosen from ureA, ureB, ureF, ureG, ureH and ureI. The
 CC invention also relates to an isolated protein encoded by the nucleic
 CC acid, a vaccine for preventing onset of disease in mammals infected by H.
 CC bizzozeronii comprising a nucleic acid and a carrier, and an isolated
 CC antibody or its binding portion raised against the nucleic acid. The
 CC nucleic acids, proteins and antibodies are useful for vaccinating mammals
 CC against onset of disease caused by infection of H. bizzozeronii, which
 CC involves administering the sequences. The sequences are useful for
 CC detecting H. bizzozeronii in a sample of tissue or body fluids which
 CC involves providing a nucleic acid as an antigen, providing an antibody,
 CC or providing a nucleotide sequence as a probe in a nucleic acid
 CC hybridisation assay, contacting the sample with the antigen or the probe,
 CC and detecting any reaction which indicates that H. bizzozeronii is
 CC present in the sample. This sequence represents the H. bizzozeronii ureB
 CC polypeptide.

XX Sequence 569 AA;

Query Match 77.6%; Score 2326; DB 8; Length 569;
 Best Local Similarity 74.1%; Pred. No. 1.2e-38;
 Matches 420; Conservative 75; Mismatches 72; Indels 0; Gaps 0;

QY 2 KMKQEVYNTYPTKGDVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMQSNPS 61
 DB 3 KISRKEYVMYPTTGDVRLGDTDLLEVEHDCTTYGEEELKFGAGKTIREGMQSNPS 62
 QY 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGKAGNKMDQGVSPHVMVGTEALA 121
 DB 63 SHELDLVLTNALIVDYTYGIYKADIGIKNGKIHGKAGNKMDQGVSPHVMVGTEALA 122
 QY 122 GEGMIITAGGIDSHTHFLSPQOFTALANGVTTFMGCGTGPVDTGNATTITPGKWNLRHM 181
 DB 123 AEGLIIVTAGGIDTHIFISPOQIPTAFASGVTIMGGTGPADGTNATTITPGKWNLRHM 182
 QY 182 LRAAEYSMNVGFLGKGNSSKKQLVEQVAGAGIKFLHEDWGTTPSAIDHCLSVADYD 241
 DB 183 LRAAEYSMNVLGFLGKGNVSEPSLVQLEAGAGIKFLHEDWGTTPSAIYHCLNVADYD 242
 QY 242 VQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTWAGELNIPSTTP 301
 DB 243 VQVAIHDTLNEAGCVEDTLQAIAGRTIHTPHTGAGGSHAPDVKMSGEFNILPASTNP 302
 QY 302 TIPTINTVAEHLDMTCHHLDRKIREDLQFSQSRIRPGSIAEDVLHDMGVIAVTSSD 361
 DB 303 TIPTVTNTEAHEMDMLVMVCHLDRKIREDLQFSQSRIRPGSIAEDVLHDMGVIAVTSSD 362
 QY 362 SOAMGRAGEVIPRTWOTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPAHGVSEY 421
 DB 363 SOAMGRAGEVITRTWQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPAHGVSEY 422
 QY 422 IGSVEEGKIADLVVNPAPFAGVKEPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMFGHH 481
 DB 423 VGSVEGKFDADLVLSAPFAGIKENMIKGGFIALSQMGDANASIPFPQPVYVREMFGHH 482
 QY 482 GKAKFDTSIITFVSKVAYENGVEKELGERQVLVKNCRNITKKDFKNDKTAKITVDPKH 541
 DB 483 GKAKFDTNITFVSVQVYDNGIKELGLQRVVLVKNCRNITKKOLKFNVDVTAHIEVNPET 542
 QY 542 FEYFVDSKLTCTSKPTSQVPLAQRVTF 568
 DB 543 YKVKVDGKEVTSKAADKISLAQLNLF 569

RESULT 7
 AAR04580
 ID AAR04580 standard; protein; 569 AA.

XX AC AAR04580;
 XX AC AAR04580;
 DT 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 21-SEP-1990 (first entry)

DE Part of protein with urease activity.

XX Urease; probe.

XX Helicobacter pylori.

PN WO9004030-A.

PD 19-APR-1990.

PF 06-OCT-1988; 88FR-00013135.

PR 06-OCT-1988; 88FR-00013135.

PA (INSP) INST PASTEUR.

PA (INRM) INSERM INST NAT SANTE & RECH MED.

XX Labigne A;

DR WPI: 1990-147844/19.

DR N-PSDB; AAQ04329.

XX New nucleotide sequences encoding Campylobacter pylori-ureasents - and
 PT derived vectors, transformants, protein, antibodies and probes, useful in
 PT diagnosis, treatment and prevention of infections.

XX Claim 12; Page 34; 47pp; French.

XX The protein can be used for the prodn. of antibodies and to prepare
 CC vaccines for the prevention /treatment of C. pylori infections. See also
 CC AAR04578-81. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24
 CC -OCT-2003 to standardise OS field)

XX Sequence 569 AA;

Query Match 76.6%; Score 2297; DB 2; Length 569;
 Best Local Similarity 73.4%; Pred. No. 4.5e-38;
 Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

QY 2 KMKQEVYNTYPTKGDVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMQSNPS 61

DB 3 KISRKEYVMYPTTGDVRLGDTDLIAEVEHDYTYGEEELKFGAGKTIREGMQSNPS 62

QY 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGKAGNKMDQGVSPHVMVGTEALA 121

DB 63 KEELDLIITNALIVDYTYGIYKADIGIKNGKIHGKAGNKMDQGVKNLNVGTEALA 122

QY 122 GEGMIITAGGIDSHTHFLSPQOFTALANGVTTFMGCGTGPVDTGNATTITPGKWNLRHM 181

DB 123 GEGLIIVTAGGIDTHIFISPOQIPTAFASGVTIMGGTGPADGTNATTITPGKWNLRHM 182

QY 182 LRAAEYSMNVGFLGKGNSSKKQLVEQVAGAGIKFLHEDWGTTPSAIDHCLSVADYD 241

DB 183 LRAAEYSMNVLGFLGKGNASNDASLADQIEAGAGIKFLHEDWGTTPSAIYHCLNVADYD 242

QY 242 VQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTWAGELNIPSTTP 301

DB 243 VQVAIHDTLNEAGCVEDTMAAIAGRTHMTFTTEGAGGSHAPDIIKVAGEHNLIPASTNP 302

QY 302 TIPTINTVAEHLDMTCHHLDRKIREDLQFSQSRIRPGSIAEDVLHDMGVIAVTSSD 361

DB 303 TIPTVTNTEAHEMDMLVMVCHLDRKIREDLQFSQSRIRPGSIAEDVLHDMGVIAVTSSD 362

Qy	362	SOAMGRAGEVIVPRTWQTADNKKKGFGLKPEDGKNDNFRIKRYISKYITNPALTHGVSEY	421
Db	363	SOAMGRVGEVITRTWQTADNKKKGFGLKKEKGNDNFRIKRYLSKYITNPALAHGISEY	422
Qy	422	IGSVBECKIADLVVNPAFCVCKPKIVIKGMVVVFSEMGDSNASVPTPOPVYVREMFEGHH	481
Db	423	VGSVEVGKADLVLSWPAFFGVKPNMIIKGGFIALSQMGDANASIPTPOPVYVREMFPAHH	482
Qy	482	GKAKFDTSTITFVSKVAYENGVKELGLERQVLPVKNCENITKDFKFNDKTAKITVDPKT	541
Db	483	GKAKYDANITFVSQAAYDKGKEELGLERQVLPVKNCENITKDMQFNDTTHAIEVNPET	542
Qy	542	FEVFDGLCTSKPTSQVPLAQRVTFP	568
Db	543	YHVFVDGKEVTSKPAKVSLAQLFSIF	569
RESULT 8			
XX	AAW07194	ID	AAW07194 standard; protein; 569 AA.
XX	AAW07194;	AC	
XX	DT	16-OCT-2003 (revised)	
XX	DT	11-FEB-1997 (first entry)	
XX	XX	H. pylori urease B subunit.	
XX	XX	Urease; ureA gene; ureB gene; vaccine.	
XX	OS	Helicobacter pylori; strain CPM630.	
XX	PN	WO9633732-A1.	
XX	PD	31-OCT-1996.	
XX	PF	25-APR-1996; 96WO-US005800.	
XX	PR	28-APR-1995; 95US-00431041.	
XX	PR	06-DEC-1995; 95US-00568122.	
XX	PA	(ORAV-) ORAVAX INC.	
XX	PI	Lee CK, Monath TP, Ackerman SK, Thomas WD, Soman G, Kleanthous H;	
XX	PI	Weltzin RA, Pappo J, Ernak T, Guirakhoo F, Bhagat H, Sussman I;	
XX	DR	WPI; 1996-497373/49.	
XX	DR	N-PSDB; AAT44351.	
XX	PT	Vaccine for inducing mucosal response to Helicobacter - contg. multimeric	
XX	PT	urease complex and pref. an antibiotic, anti-secretory agent or bismuth	
XX	PS	salt.	
XX	PS	Disclosure; Page 71-73; 98pp; English.	
XX	CC	Urease A (AAW07193) and B (AAW07194) subunits of Helicobacter pylori are	
XX	CC	encoded by the ureA + ureB gene locus of clinical isolate CPM630.	
XX	CC	Vectors, e.g. pORV214 (see also AAT44351), carrying the gene locus can be	
XX	CC	used to transform host cells for the large-scale prodn. of recombinant,	
XX	CC	enzymatically inactive, multimeric urease. The urease complex induces a	
XX	CC	mucosal immune response that can treat or prevent Helicobacter, esp. H.	
XX	CC	pylori, gastroduodenal infection. (Updated on 16-OCT-2003 to standardise	
XX	CC	OS field)	
XX	SQ	Sequence 569 AA;	
Query Match 76.6%; Score 2297; DB 2; Length 569;			
Best Local Similarity 73.4%; Pred. No. 4.5e-38;			
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;			
Qy	2	KMKKQEVYNTVPTGDKVRLGDTDLMAEVBHDYTYTGEELKFGAGKTIREGMGQSNPSD	61
Db	3	KIRSRKEVSMYGPPTGDKVRLGDTDLMAEVBHDYTYTGEELKFGGCKTLREGMSQSNPS	62

Qy	62	ENTLDLVITNAMILDYTGIVKADIGIKNKGKIHGKAGNKMQDGVSPHMVVGVGTEALA	121
Db	63	KEELDLDIITNALIVDYTGIVKADIGIKDGKTAGIKGKGNKMQDGVKNNLSVGPATEALA	122
Qy	122	GGEMITITAGGIDSHTHFLSPQOPFTALANGVTTWFGGSGTGPVDGNTATTITPGKNNLHRM	181
Db	123	GGELVITAGGIDTHIHFIISQOIIPTAFASGVTTMIGGSGTPADGNTATTIIPGRNRLKWM	182
Qy	182	LRAAEYSMNVGFLGKGNSSKKOLVQVEAGAIGFKLHDWGTTTPSAIDHCLSVADEYD	241
Db	183	LRAAEYSMNLGFLAKGNASNDASLADQIEAGAIGFKLHDWGTTTPSAIHNALDVADKYD	242
Qy	242	VQVCIHDTTVNEAGYDDTLNANNGRAIHAYHIEGAGGCHSPDVTIMAGELNILPSSITTP	301
Db	243	VQVAIHDTTLNEAGCVEDTWAAYIAGRTWHTFHTTEGAGGCHADPIIKVAGEHNILPASTNP	302
Qy	302	TIPVTINTVAEHLDMLTCHHLDKRISLEDLOFSOSRIRPGSIAAEDVLHDMGVIAMTSSD	361
Db	303	TIPFTVNTAEAHMDMLVCHHLDSIKEDVOFADSRIRPQTIAAEDTLHDMGIFSIITSD	362
Qy	362	SQAMGRAGEVIPRTWQTADKNKBFGLKPEDGKDNDFRIKRYISKYITINPALTHGVSEY	421
Db	363	SQAMGRGVEVITRTWQTADKNKBFGLKKEBKGDNDNFRIKRYLSKYITINPAIAHGISEY	422
Qy	422	IGSVESGKIADLVVNNPAFGVCKPIVIKGMVVFSEMGDSNASVPTPQPVYVYREMGFHH	481
Db	423	VGSVEVGKVADLVWSFAFGVCKPNMLIKGGFIALSQMGDANASIPTPQPVYVYREMFHH	482
Qy	482	GKAKFDTSIITFVSKVAVYENGKELGLERQVLVKNKCNITKQDFKPNDKTAKITVDPKPT	541
Db	483	GKAKYDANIITFVSQAADKGIKEELGLERQVLVKNKCNITKQDFNDTTHAIEVNPET	542
Qy	542	FEYFVVDGKLTCKPTSQVPIAQRVTFP	568
Db	543	YHFVVDGKEVTSKPANKVSLAQLFSIF	569
RESULT 9			
AAW98511			
XX	ID	AAW98511 standard; protein; 569 AA.	
AC	AAW98511;		
XX			
DT	31-MAR-1999	(first entry)	
XX			
DE	H. pylori	GHPO 1248 protein.	
KW		GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;	
KW		peptic ulcer disease.	
XX			
OS		Helicobacter pylori.	
XX			
PN	WO9843478-A1.		
XX			
PD	08-OCT-1998.		
XX			
PF	01-APR-1998;	98WO-US006371.	
XX			
PR	01-APR-1997;	97US-00833457.	
PR	24-JUN-1997;	97US-00881227.	
PR	29-JUL-1997;	97US-00902615.	
XX			
PA	(INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.		
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
XX	Kleanthous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;		
PI			
XX			
XX	WPI; 1998-542293/46.		
DR	N-PSDB; AAX14230.		
XX			
PT	New isolated Helicobacter polynucleotides - used to develop products for the diagnosis, prevention and treatment of Helicobacter infections and		

New isolated *Helicobacter* polynucleotides - used to develop products for the diagnosis, prevention and treatment of *Helicobacter* infections and


```

Db 123 GEGLVITAGGIDTHIHFIPOQITAFASGVTTMIGGTGPDGNTATTTPGRNLKWM 182
QY 182 LRAAEYSNMVGFLLKGNSSKKQLVEQVAGAIQFKLHEDWGTTPSAIDHCLSVADYD 241
Db 183 LRAAEYSNMVGFLLKGNSSNDASLADQIEAGAIQFKLHEDWGTTPSAINHALDVADYD 242
QY 242 VQVCIHTDTNEAGYVDDTLNMGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSSTP 301
Db 243 VQVAIHDTTNEAGCVEDTMAAIAGRWTHTFHTGAGGSHAPDIIVKAGEHNILPASTNP 302
QY 302 TIPVTINTVAEHLDMVCHHLDKRIEDLOFSQSRIRPGSIAAEDVLHDMGVIAMTSSD 361
Db 303 TIPVTNTEAEHMDMLVCHHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSIYSSD 362
QY 362 SQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKDNDFRIKRYISKYITINPALTHGVSEY 421
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGKLPEDGKDNDFRIKRYISKYITINPALTHGVSEY 422
QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFHH 481
Db 423 VGSVEVGKADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFHH 482
QY 482 GKAKPDSITFVSKVAYENGVEKLGRLERQVLPVKNCNITKKDPKNDKTAKITVDPKT 541
Db 483 GKAKYDANITFVSOAAYDKGIEKELGLERQVLPVKNCNITKKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKCLTSKPTSOVPLAQRVTF 568
Db 543 YHVFVDGKETSVPANKVSLAQLFSIP 569

```

```

RESULT 11
ID ADU05536 standard; protein; 569 AA.
XX AC ADU05536;
XX DT 27-JAN-2005 (first entry)
XX DE H. pylori antigenic protein HP0072.
XX KW antibacterial; antigenic; H. pylori infection.
XX OS Helicobacter pylori.
XX PH Key Location/Qualifiers
XX FT 533..567
XX FT /note= "Immunogenic region. Region specifically claimed
XX FT in claim 14"
XX PN WO2004094467-A2.
XX PD 04-NOV-2004.
XX PF 22-APR-2004; 2004WO-EP004255.
XX PR 22-APR-2003; 2003EP-00450097.
XX PA (INTE-) INTERCELL AG.
XX PI Meinke A, Min Bui D, Nagy E, Henics T;
XX DR WPI; 2004-775908/76.
XX DR N-PSDB; ADU05358.
XX PT New hyperimmune serum reactive antigens from Helicobacter pylori, and
XX PT encoding nucleic acid molecules, useful for diagnosing, preventing or
XX PT treating H. pylori infections.
XX PS Claim 14; SEQ ID NO 184; 176pp; English.
XX CC The invention relates to an isolated nucleic acid molecule encoding a

```

```

CC hyperimmune serum reactive antigen or its fragment. The composition
CC (including the nucleic acid molecule, hyperimmune serum-reactive antigen
CC or antibody) is useful for manufacturing a medicament or pharmaceutical
CC preparations (e.g. a vaccine) for treating or preventing H. pylori
CC infections. The antigen or its fragment may also be used for isolating,
CC purifying and/or identifying an interaction partner of the hyperimmune
CC serum reactive antigen or fragment; for generating a peptide binding to
CC the hyperimmune serum reactive antigen or fragment, where the peptide is
CC selected from anticalins; for manufacturing a functional nucleic acid
CC selected from aptamers and spiegelmers; or for manufacturing a functional
CC ribonucleic acid selected from ribozymes, antisense nucleic acids and
CC siRNA. The present sequence represents the amino acid sequence of an H.
CC pylori antigenic protein.
XX SQ Sequence 569 AA;

Query Match 76.6%; Score 2297; DB 8; Length 569;
Best Local Similarity 73.4%; Pred. No. 4.5e-38;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

QY 2 KMKQEQYNTYGTPTKDGKVRGLGDTDLMAEVEHDYTYGELKFGAGKTIREGMGQSNSPD 61
Db 3 KISRKEYVSMYGTPTGDKVRLGDTDLIAEVEHDYTYGELAFGGKTLREGMSQSNPS 62
QY 62 ENTLDLVTNAMIIDYTGIIYKADIGIKNGKIHGIGKAGKMDQGVSPHMVVGVTGTEALA 121
Db 63 KEELDLIITNALIVDYGIIYKADIGIKDGKIAGIKGGKMDQGVKNNLSVGPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQOFTALANGVTTMFGGGTGPVDGNTATTITPGKNNLHRM 181
Db 123 GEGLVITAGGIDTHIHFIPOQITAFASGVTTMIGGTGPDGNTATTITPGRNLKWM 182
QY 182 LRAAEYSNMVGFLLKGNSSKKQLVEQVAGAIQFKLHEDWGTTPSAIDHCLSVADYD 241
Db 183 LRAAEYSNMVGFLLKGNSSNDASLADQIEAGAIQFKLHEDWGTTPSAINHALDVADYD 242
QY 242 VQVCIHTDTNEAGYVDDTLNMGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSSTP 301
Db 243 VQVAIHDTTNEAGCVEDTMAAIAGRWTHTFHTGAGGSHAPDIIVKAGEHNILPASTNP 302
QY 302 TIPVTINTVAEHLDMVCHHLDKRIEDLOFSQSRIRPGSIAAEDVLHDMGVIAMTSSD 361
Db 303 TIPVTNTEAEHMDMLVCHHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSIYSSD 362
QY 362 SQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKDNDFRIKRYISKYITINPALTHGVSEY 421
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGKLPEDGKDNDFRIKRYISKYITINPALTHGVSEY 422
QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFHH 481
Db 423 VGSVEVGKADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFHH 482
QY 482 GKAKPDSITFVSKVAYENGVEKLGRLERQVLPVKNCNITKKDPKNDKTAKITVDPKT 541
Db 483 GKAKYDANITFVSOAAYDKGIEKELGLERQVLPVKNCNITKKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKCLTSKPTSOVPLAQRVTF 568
Db 543 YHVFVDGKETSVPANKVSLAQLFSIP 569

RESULT 12
ADU09177
ID ADS09177 standard; protein; 568 AA.
XX AC ADS09177;
XX DT 16-DEC-2004 (first entry)
XX DE H. pylori urease beta subunit.
XX KW Urease; beta subunit; UreB; gastric ulcer; peptic ulcer; gastritis;
KW stomach cancer; vaccine; antibody; immune reaction.

```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:43:17 ; Search time 51 Seconds
(without alignments)
974.851 Million cell updates/sec

Title: US-09-904-994B-3
Perfect score: 2399
Sequence: 1 MKMKQBYVNTYGTGDKV.....KLCTSKPTSQVPLAQRVTF 568

Scoring table: BLOSUM62
Gapop 4.0 , Gapext 1.0

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC Cellerai_SIDS3/prodata/2/iaa/5_COMB.pap.*
- 2: /EMC Cellerai_SIDS3/prodata/2/iaa/6_COMB.pap.*
- 3: /EMC Cellerai_SIDS3/prodata/2/iaa/7_COMB.pap.*
- 4: /EMC Cellerai_SIDS3/prodata/2/iaa/H_COMB.pap.*
- 5: /EMC Cellerai_SIDS3/prodata/2/iaa/PCTUS_COMB.pap.*
- 6: /EMC Cellerai_SIDS3/prodata/2/iaa/RE_COMB.pap.*
- 7: /EMC Cellerai_SIDS3/prodata/2/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2304	76.8	568	2	US-09-338-920B-12
2	2297	76.6	569	2	US-09-338-920B-6
3	2294	76.5	569	1	US-08-467-822-26
4	2294	76.5	569	2	US-08-432-697-26
5	2294	76.5	569	2	US-08-466-248-26
6	2290	76.4	566	1	US-08-920-095-3
7	2290	76.4	566	5	PCT-US96-05800-3
8	2282	76.1	570	2	US-09-431-705-5
9	2284	74.8	569	1	US-08-467-822-21
10	2244	74.8	569	2	US-08-432-697-21
11	2244	74.8	569	2	US-08-466-248-21
12	2244	74.8	569	2	US-09-338-920B-10
13	2237	74.6	559	2	US-09-338-920B-8
14	2044	68.2	568	1	US-07-732-242C-3
15	1834	61.2	566	2	US-09-252-991A-26887
16	1831	61.1	569	1	US-08-457-822-27
17	1831	61.1	569	2	US-08-432-697-27
18	1831	61.1	569	2	US-08-466-248-27
19	1831	61.1	569	2	US-09-543-681A-6029
20	1830	61.0	625	2	US-09-489-039A-9216
21	1829	61.0	567	2	US-09-328-352-5912
22	1826	60.9	567	1	US-08-967-513-5
23	1826	60.9	567	1	US-08-687-645B-5
24	1804	60.2	571	2	US-09-710-279-118
25	1804	60.2	840	1	US-08-467-822-25
26	1804	60.2	840	2	US-08-432-697-25

ALIGNMENTS

RESULT 1

US-09-338-920B-12
; Sequence 12, Application US/09338920B
; Patent No. 6709851

GENERAL INFORMATION:

; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter heilmannii

Query Match	76.8%	Score 2304;	DB 2;	Length 568;
Best Local Similarity	73.7%;	Pred. No. 1e-40;		
Matches 418;	Conservative 74;	Mismatches 74;	Indels 1;	Gaps 1;
Qy	2	MMKKQBYVNTYGTGDKVRLGDTDLWAEVHDYTYGELKFGAGKTIREGMGQNSPD	61	
Db	3	KISREYVSMVGTGDKVRLGDTDLLEVEHDTTYGEEIKFGGKTIRDGNGQINS	62	
Qy	62	ENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHIKAGNKMDDGVSPHMVVGTEALA	121	
Db	63	SHELDLVTNALTIVDVTGIYKADIGIKNGKIHIKAGNKLQDQVGNLVCVGPATEALA	122	
Qy	122	GEGLMIITAGGDSHTHFLSPQFPQFALANGVTMTGGTGPVDGTNATITPGKWLHRM	181	
Db	123	AEGLVTAGGIDTHIFHSPOQITAFASGITMTGGTGPADGTNATITPGRWLKEK	182	
Qy	182	LRAAEYSNVGFLKGNSSKKQLVEQVAGAIQFKLHEDMGTTPTSAIDHCLSVADEYD	241	
Db	183	LRASEYANNLGYLGNVSPFALIDQEAQAIQFKIHEDMGSTFSAINHALNIADKYD	242	
Qy	242	VQVCITHTVNEAGYVDDTLNANNGRAIHAYHIEGAGGSHSPDVTIMAGELNIPSTTP	301	
Db	243	VQVALHTDTLNEAGCVEDTLEAIGRTIHTFTEGAGGGHAPDVIKWAGEFNILPASTNP	302	
Qy	302	TIPYINTVAEHLDMTCHHLDKKIREDLQFSQSRIRPGSTAAEDVLHDMGVIAWTSSD	361	
Db	303	TIPFTKNTAEAHMDLM-CHHLDKNIKEVFEADSRIRPQTIAAEDKLHDMGIFITS	361	

QY 362 SOAMGRAGEVLPRTWQTADKMKKBEKGLPEKGDNDNPNRIKRYISKYKTYINPALTHGVSEY 421
 Db 362 SOAMGRAGEVLPRTWQTADKMKKBEKGLPEKGDNDNPNRIKRYISKYKTYINPALTHGVSEY 421
 QY 422 IGSVEEGKIADLVVNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVVYREMFHGH 481
 Db 422 VGSVEGKIADLVVNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVVYREMFHGH 481
 QY 482 GKAKFDTSITFVSKVAYENGVEKGLERQVLVVKNCRNITKQPKFNDKTAKITVDPKT 541
 Db 482 GKAKFDNITFVSKVAYENGVEKGLERQVLVVKNCRNITKQPKFNDKTAKITVDPKT 541
 QY 542 FEVVDGKLTCKTSKPTSQVPLAQRITFF 569
 Db 542 YKVKVDGNEVTSHAADKLSLAQLNLF 568
 RESULT 2
 US-09-338-920B-6
 ; Sequence 6, Application US/09338920B
 ; Patent No. 6709851
 ; GENERAL INFORMATION:
 ; APPLICANT: Soman, Gopalan
 ; APPLICANT: Thomas, William D., Jr.
 ; TITLE OF INVENTION: Stabilization of Helicobacter Urease
 ; FILE REFERENCE: 06132/023002
 ; CURRENT APPLICATION NUMBER: US/09/338,920B
 ; CURRENT FILING DATE: 2002-12-23
 ; PRIOR APPLICATION NUMBER: US 08/928,081
 ; PRIOR FILING DATE: 1997-09-12
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 569
 ; TYPE: PRT
 ; ORGANISM: Helicobacter pylori
 US-09-338-920B-6
 Query Match 76.6%; Score 2297; DB 2; Length 569;
 Best Local Similarity 73.4%; Pred. No. 1.4e-40;
 Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
 QY 2 KMKQEVNTYGTGKDKVRLGDTDLMAVEHDYTYTGEEELKFGAGKTIREGMSQNSPD 61
 Db 3 KISRKEVSMVPTGDKVRLGDTDLMAVEHDYTYTGEEELKFGAGKTIREGMSQNSPD 62
 QY 62 ENTLDLVITNAMIIDYTIYKADIGIKNGKIHGKAGKQMDQGVSPHVVGVGTAL 121
 Db 63 KEELDLITNALIVDYTIYKADIGIKNGKIHGKAGKQMDQGVSPHVVGVGTAL 122
 QY 122 GEGMIITAGGIDSHFSLPOPTALANGVTTFEGGTGPVDTGNATITPGKWLHRM 181
 Db 123 GEGMIITAGGIDSHFSLPOPTALANGVTTFEGGTGPVDTGNATITPGKWLHRM 182
 QY 182 LRAAEYSVMNVGFLKGNSSKOLVQVEAGAIGFLKHEHDWGTTPSAIDHCLSLVADEVD 241
 Db 183 LRAAEYSVMNVGFLKGNSSKOLVQVEAGAIGFLKHEHDWGTTPSAIDHCLSLVADEVD 242
 QY 242 VQVCITDVTNEAGVDDTLNAMGRAIHAYHIEGAGGHSFDVITMAGELNIPSSSTP 301
 Db 243 VQVAIHTDNEAGVDDTLNAMGRAIHAYHIEGAGGHSFDVITMAGELNIPSSSTP 302
 QY 302 TIPTVNTVAEHLDMTCHHLDKRIRDLQFSQRIIPGSIADVEDVLDGMVIAWTSDD 361
 Db 303 TIPTVNTVAEHLDMTCHHLDKRIRDLQFSQRIIPGSIADVEDVLDGMVIAWTSDD 362
 QY 362 SOAMGRAGEVLPRTWQTADKMKKBEKGLPEKGDNDNPNRIKRYISKYKTYINPALTHGVSEY 421
 Db 363 SOAMGRAGEVLPRTWQTADKMKKBEKGLPEKGDNDNPNRIKRYISKYKTYINPALTHGVSEY 422
 QY 422 IGSVEEGKIADLVVNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVVYREMFHGH 481

Db 423 VGSVEGKVADLVVNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVVYREMFHGH 482
 QY 482 GKAKFDTSITFVSKVAYENGVEKGLERQVLVVKNCRNITKQPKFNDKTAKITVDPKT 541
 Db 483 GKAKYDANITFVSKVAYENGVEKGLERQVLVVKNCRNITKQPKFNDKTAKITVDPKT 542
 QY 542 FEVVDGKLTCKTSKPTSQVPLAQRITFF 568
 Db 543 YHVFVDGKEVTSKPKVKSQAQLFSIF 569
 RESULT 3
 US-08-467-822-26
 ; Sequence 26, Application US/08467822
 ; Patent No. 5843460
 ; GENERAL INFORMATION:
 ; APPLICANT: Labigne, Agnes
 ; APPLICANT: Sauerbaum, Sebastien
 ; APPLICANT: Ferrero, Richard L.
 ; APPLICANT: Thiherge, Jean-Michel
 ; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
 ; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
 ; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
 ; NUMBER OF SEQUENCES: 44
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 ; ADDRESS: Dunner
 ; STREET: 1300 I Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/467,822
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/447,177
 ; FILING DATE: 19-MAY-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/432,697
 ; FILING DATE: 02-MAY-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meyers, Kenneth J.
 ; REGISTRATION NUMBER: 25,146
 ; REFERENCE/DOCKET NUMBER: 03495.0137-02000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 408-4000
 ; TELEFAX: (202) 408-4400
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 569 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-467-822-26
 Query Match 76.5%; Score 2294; DB 1; Length 569;
 Best Local Similarity 73.4%; Pred. No. 1.6e-40;
 Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
 QY 2 KMKQEVNTYGTGKDKVRLGDTDLMAVEHDYTYTGEEELKFGAGKTIREGMSQNSPD 61
 Db 3 KISRKEVSMVPTGDKVRLGDTDLMAVEHDYTYTGEEELKFGAGKTIREGMSQNSPD 62

```
Db 3 KISRKEYVMVPTGDKVRLGDTDLIAEVEHDYTYIGBELKFGGKTLREGMSQSNPNS 62
Qy 62 ENTLDLVTNAMIIDYTYIKADIGIKNGKIHGKAGKMDQGVSPHVVGVGTAL 121
Db 63 KEELDLIITNALIVDTYTYIKADIGIKNGKIHGKAGKMDQGVKNLSVGPATEALA 122
Qy 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTFGGGTGPVDTGNATTITPGKNLHRM 181
Db 123 GEGLIIVTAGGIDTHIFISPOQIPAPASGVTTMIGGGTGPADGTNATTITPGRRNLKWM 182
Qy 182 LRAAEYSNMVGLFKGNSSKKQLVEQVAGAIKFKLHEDMGTTTSAIDHCLSVAD 241
Db 183 LRAAEYSNMVGLFKGNSSKKQLVEQVAGAIKFKLHEDMGTTTSAIDHCLSVAD 242
Qy 242 VQVCHTDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGCHSPDVTMAGELNIPSS 301
Db 243 VQVAIHTDTLNEAGCVEDTMAAIAAGTMTFTHTGAGGCHAPDIIKVAGEHNILPASTNP 302
Qy 302 TIPTVTINTVAEHLDMMLTCHLDKRIREDLQPSQSRIRPGSIAAEDVLHDMGVIA 361
Db 303 TIPTVTINTVAEHLDMMLTCHLDKRIREDLQPSQSRIRPGSIAAEDVLHDMGVIA 362
Qy 362 SOAMGRAGEVITRTWTADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSEY 421
Db 363 SOAMGRAGEVITRTWTADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSEY 422
Qy 422 IGSVEEGKIADLVNMPAFFGKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFGHH 481
Db 423 VGSVEVGKVDLVNMPAFFGKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFGHH 482
Qy 482 GKAKFDTSTITFVSKVAYENGVEKELGLERQVLVKNCRNITKKDFKNDKTAKITVDPKT 541
Db 483 GKAKYDRNITFVSOAAYDKGKEELGLERQVLVKNCRNITKKDMQFNDTTAIEVNPET 542
Qy 542 FEVFDGKLTCKTSKPTSOVPLAQRTPF 568
Db 543 YHVFVDGKVTSPANKVSLAQLFSIF 569

RESULT 4
US-08-432-697-26
; Sequence 26, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432.697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
```

```
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-432-697-26

Query Match 76.5%; Score 2294; DB 2; Length 569;
Best Local Similarity 73.4%; Pred. No. 1.6e-40;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

Qy 2 KMKKQRYVNTYPTGDKVRLGDTDLIAEVEHDYTYIGBELKFGGKTLREGMSQSNPNS 61
Db 3 KISRKEYVMVPTGDKVRLGDTDLIAEVEHDYTYIGBELKFGGKTLREGMSQSNPNS 62
Qy 62 ENTLDLVTNAMIIDYTYIKADIGIKNGKIHGKAGKMDQGVSPHVVGVGTAL 121
Db 63 KEELDLIITNALIVDTYTYIKADIGIKNGKIHGKAGKMDQGVKNLSVGPATEALA 122
Qy 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTFGGGTGPVDTGNATTITPGKNLHRM 181
Db 123 GEGLIIVTAGGIDTHIFISPOQIPAPASGVTTMIGGGTGPADGTNATTITPGRRNLKWM 182
Qy 182 LRAAEYSNMVGLFKGNSSKKQLVEQVAGAIKFKLHEDMGTTTSAIDHCLSVAD 241
Db 183 LRAAEYSNMVGLFKGNSSKKQLVEQVAGAIKFKLHEDMGTTTSAIDHCLSVAD 242
Qy 242 VQVCHTDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGCHSPDVTMAGELNIPSS 301
Db 243 VQVAIHTDTLNEAGCVEDTMAAIAAGTMTFTHTGAGGCHAPDIIKVAGEHNILPASTNP 302
Qy 302 TIPTVTINTVAEHLDMMLTCHLDKRIREDLQPSQSRIRPGSIAAEDVLHDMGVIA 361
Db 303 TIPTVTINTVAEHLDMMLTCHLDKRIREDLQPSQSRIRPGSIAAEDVLHDMGVIA 362
Qy 362 SOAMGRAGEVITRTWTADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSEY 421
Db 363 SOAMGRAGEVITRTWTADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSEY 422
Qy 422 IGSVEEGKIADLVNMPAFFGKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFGHH 481
Db 423 VGSVEVGKVDLVNMPAFFGKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFGHH 482
Qy 482 GKAKFDTSTITFVSKVAYENGVEKELGLERQVLVKNCRNITKKDFKNDKTAKITVDPKT 541
Db 483 GKAKYDRNITFVSOAAYDKGKEELGLERQVLVKNCRNITKKDMQFNDTTAIEVNPET 542
Qy 542 FEVFDGKLTCKTSKPTSOVPLAQRTPF 568
Db 543 YHVFVDGKVTSPANKVSLAQLFSIF 569

RESULT 5
US-08-466-248-26
; Sequence 26, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
```


;; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
;; ADDRESSEE: Dunner
;; STREET: 1300 I Street, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005-3315
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/466,248
;; FILING DATE: 06-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/447,177
;; FILING DATE: 19-MAY-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/432,697
;; FILING DATE: 02-MAY-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meyers, Kenneth J.
;; REGISTRATION NUMBER: 25,146
;; REFERENCE/DOCKET NUMBER: 03495.0137-02000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 408-4000
;; TELEFAX: (202) 408-4400
;; INFORMATION FOR SEQ ID NO: 26:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 569 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-466-248-26

Query Match 76.5%; Score 2294; DB 2; Length 569;
Best Local Similarity 73.4%; Pred. No. 1.6e-40;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
QY 2 KMKQBYVNTYPTGDKVRLGDTDLAEVHEHYTYGEEKLFGAGKTIREGMSQNSPD 61
DB 3 KISRKEYVSMYPTGDKVRLGDTDLAEVHEHYTYGEEKLFGAGKTIREGMSQNSPD 62
QY 62 ENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDDGVSPHMVGVGTEALA 121
DB 63 KEELDLIITNALIVDITYGIYKADIGIKNGKIHGIGKAGNKMDDGVSPHMVGVGTEALA 122
QY 122 GEGMIITAGGIDSHFLSPQFPPTALANGVTTFMGGGTGPDVGTNATTITPGKWNLRHM 181
DB 123 GEGLVTAGGIDTHIFISPOQIPTAFASGVTTMIGGGTGPADGTNATTITPGKWNLRHM 182
QY 182 LRAAEYSNMVGFLLGKNSKKQLVEOVBAIGIKNGKIHGIGKAGNKMDDGVSPHMVGVGTEALA 241
DB 183 LRAAEYSNMVGFLLGKNSKKQLVEOVBAIGIKNGKIHGIGKAGNKMDDGVSPHMVGVGTEALA 242
QY 242 VQVCITDVTNAGVDDTLNMGRAITHAYHIEGAGGSHSPDVTITMAGELNLPSSSTP 301
DB 243 VOVAHTDTLNEAGVEDTMAIAGRTWHTFHTGAGGGHAPDIIVKAGEHNILPASTNP 302
QY 302 TIPTVINTVAHLDMKMTCHHLDRKREDLQFSQSRIRPGSIAEDVLHDMGVIAWTSDD 361
DB 303 TIPTVNTAEASHMDMLVACHHLDRKREDLQFSQSRIRPGSIAEDVLHDMGVIAWTSDD 362
QY 362 SOACRAGEVTPRTWQTADKNGKFGKLPEDGKONDNERIKRYISKYITINPALTHGVSEY 421
DB 363 SOACRAGEVTPRTWQTADKNGKFGRLKBEKGDNDFRIKRYLSKYITINPAIAHGYSEY 422
QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGWVFPSEMGDSNASVPTPQPVVYREMFHGH 481

DB 423 VGSVEGVKADLVLWSPAPFGVKPNMIRKGGFIALSQMGDANASIPTPQPVVYREMFHGH 482
QY 482 GKAKEDTSTITFVSKVAYENGVEKGLERQVLFPVKNCRNITKKDKFNDKTKAKITVDPKT 541
DB 483 GKAKYDRNITFVSQAAYDKGKEELGLERQVLFPVKNCRNITKKDKFNDKTKAKITVDPKT 542
QY 542 FEVFDGKLCSTKPTSQVPLAQRYTFF 568
DB 543 YHVFVDGKEVTSKPNKVSQAQLFSIF 569
RESULT 6
US-08-920-095-3
; Sequence 3, Application US/08920095
; Patent No. 5837240
; GENERAL INFORMATION:
; APPLICANT: Cynthia K. Lee et al.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,095
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,041
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-920-095-3

Query Match 76.4%; Score 2290; DB 1; Length 566;
Best Local Similarity 73.6%; Pred. No. 2e-40;
Matches 415; Conservative 75; Mismatches 74; Indels 0; Gaps 0;
QY 2 KMKQBYVNTYPTGDKVRLGDTDLAEVHEHYTYGEEKLFGAGKTIREGMSQNSPD 61
DB 3 KISRKEYVSMYPTGDKVRLGDTDLAEVHEHYTYGEEKLFGAGKTIREGMSQNSPD 62
QY 62 ENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDDGVSPHMVGVGTEALA 121
DB 63 KEELDLIITNALIVDITYGIYKADIGIKNGKIHGIGKAGNKMDDGVSPHMVGVGTEALA 122
QY 122 GEGMIITAGGIDSHFLSPQFPPTALANGVTTFMGGGTGPDVGTNATTITPGKWNLRHM 181
DB 123 GEGLVTAGGIDTHIFISPOQIPTAFASGVTTMIGGGTGPADGTNATTITPGKWNLRHM 182
QY 182 LRAAEYSNMVGFLLGKNSKKQLVEOVBAIGIKNGKIHGIGKAGNKMDDGVSPHMVGVGTEALA 241
DB 183 LRAAEYSNMVGFLLGKNSKKQLVEOVBAIGIKNGKIHGIGKAGNKMDDGVSPHMVGVGTEALA 242

[illegible]

RESULT 7
PCT-US96-05800-3
Sequence 3, Application PC/TUS9605800
GENERAL INFORMATION:
APPLICANT: OraVax, Inc.
TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05800
FILING DATE: 23-APR-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,041
FILING DATE: 28-APR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/568,122
FILING DATE: 06-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06132/020001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 566 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-05800-3

```

Best Local Similarity   73.6%; Pred. No. 2e-40;
Matches 415; Conservative 75; Mismatches 74; Indels 0; Gaps 0

QY      2 KMKKEVNTYGPPTGDKVRIGDITDLWAEVEVDHYTTTGEELKFGAGKTIREGMGOSNPSD 61
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db      3 KISKSEVSMVGPTTGDKVRIGDITDLTAEBEDHTTYIGEELKFGGKTIRMGSSQNPS 62
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
QY     62 ENTLDLVTNAMIIDTYGIYKADIGIKNGKIHGIGKAGNKMDQGVSPHMVVGVGTEALA 121
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db     63 KEELDIIITNALIVDYTYGIYKADIGIKDGKIAGIKGKNKDQGVVKNNLSVGPATEALA 122
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
QY    122 GEGMIITAGGISDTHFLSPPOFTALANGVTTTFGGGTGPGVDGNATTTIIPGKWNLRM 181
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db    123 GEGLIIVTAGGITDIHIFISPOOIPTAFASGVYTTMIGGGTGPDAGTNATTITPGRRLKWM 182
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
QY    182 LRAABEYSMVNGFTLGKNGSSKKOLVQVEAGAIGFKIHEDWGTTTPSAIDHCLSVADEVD 241
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db    183 LRAABEYSMNLGFLAKNGASNDSLADOIEAAGIFKIHEDWGTTTPSAINHALDVADKYD 242
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
QY    242 VQVCIHDTVNEAGYVDITLANMNGRAIHAYHIEGAGGHGSPDVTIWAGELNILPSSTP 301
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db    243 VQVAIHDTLINEACVEDTWAAIAGRTWHIFTETGAGGHPADIIVKAGEHNILPASTNP 302
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
QY    302 TIPTYINTVAEHLDMLTCHHLDKRIBEDLQFSORIRPGSIAAADVLHDWGVIAMTSSD 361
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db    303 TIPPVTNTEAHMDMLVCHHLDKSKIEDVQFADSRIRPOTIAAEDTLHDWGIPTSITSSD 362
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
QY    362 SQAMGRAGEVIPRTWTQADRNKKFEKGLPEDCKNDNFRIKRYISKYTINPALTHGVSEY 421
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db    363 SQAMGRVGEVITRTWTQADRNKKFEKGLKEEKGDNDNFRIKRYLSKYTINPAIHAHISEY 422
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
QY    422 IGSVEECKIADLVVNPNFAFFGVKPVIKGMVMVFSEMGSDNASVPPTQPYYREMFGHH 481
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db    423 VGSVEGKVADLVLSPAFFGVKPFNMIIKGFIALSQMGDNASIPPTQPYYREMFAHH 482
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
QY    482 GKAKFDTSITFVSKVAYENGVEKGLGLERQVLVYKNCRNITTKDFKFNDAKTAKITVDPKT 541
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db    483 GKAKYDANIITFVSQAAYDKGKEELGLERQVLVYKNCRNITTKDQFNDDTTAHIEVNET 542
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
QY    542 FEVFDVGKLCTSKPTSQVPLAQRY 565
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db    543 YHVFDVGKEVTSKPANKVSLAQLF 566
       : :::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|

RESULT 8
US-09-431-705-5
; Sequence 5, Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Londono-Arcila, Patricia
; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; ; TITLE OF INVENTION: vaccination against helicobacter infection
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431,705
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-431-705-5

```

```

Query Match      76.1%; Score 2282; DB 2; Length 570;
Best Local Similarity 73.0%; Pred. No. 2.9e-40;
Matches 416; Conservative 75; Mismatches 76; Indels 3; Gaps 1;

Qy 2 KMKK---QYVNTYPTGDKVRLGDTDLAAVEHDYTTYGEELKFGAGKTIREGMQSN 58
   |||||
Db 1 EMKKISRKEVSMYPTGDKVRLGDTDLAAVEHDYTTYGEELKFGGGKTLREGMQSN 60

```

Query Match 76.4%; Score 2290; DB 5; Length 566;

QY	362	SOAMGRAGEVITRTWQTADKNKKEFGKLPEDCKNDNPFRIKYISKYITINPALTGHVSEY	421
DB	363	SOAMGRVGEVITRTWQTADKNKKEFGKLPEDCKNDNPFRIKYISKYITINPALTGHVSEY	422
QY	422	IGSVEEGKADLVVNPPAFPGVKPKIVIKGMMVSEMGDSNASVPTPOPVYVREMFHGH	481
DB	423	GVSEVGYKADLVVNPPAFPGVKPKIVIKGMMVSEMGDSNASVPTPOPVYVREMFHGH	482
QY	482	GKAKPDTSTITFVSKVAYENGKVKELGLERQVLFPVKNCRNITKKPKFNDKTAKITVDPKT	541
DB	483	GKAKPDTSTITFVSKVAYENGKVKELGLERQVLFPVKNCRNITKKPKFNDKTAKITVDPKT	542
QY	542	FEVFDGKLTCKTSKPTSQVPLAQRYTTP	568
DB	543	YKVKVDGKVTSKADELSLAQLYNLF	569
<p>RESULT 11</p> <p>US-08-466-248-21</p> <p>Sequence 21, Application US/08466248</p> <p>Patent No. 6258359</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Labigne, Agnes</p> <p>APPLICANT: Sauerbaum, Sebastien</p> <p>APPLICANT: Ferrero, Richard L.</p> <p>APPLICANT: Thiberge, Jean-Michel</p> <p>TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST</p> <p>TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE</p> <p>TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID</p> <p>TITLE OF INVENTION: POLYPEPTIDES</p> <p>NUMBER OF SEQUENCES: 44</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &</p> <p>ADDRESSEE: Dunner</p> <p>STREET: 1300 I Street, N.W.</p> <p>CITY: Washington</p> <p>STATE: D.C.</p> <p>COUNTRY: USA</p> <p>ZIP: 20005-3315</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patent In Release #1.0, Version #1.30</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/466,248</p> <p>FILING DATE: 06-JUN-1995</p> <p>CLASSIFICATION: 435</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: US 08/447,177</p> <p>FILING DATE: 19-MAY-1995</p> <p>CLASSIFICATION: 435</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: US 08/432,697</p> <p>FILING DATE: 02-MAY-1995</p> <p>CLASSIFICATION: 435</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Meyers, Kenneth J.</p> <p>REGISTRATION NUMBER: 25,146</p> <p>REFERENCE/DOCKET NUMBER: 03495.0137-02000</p> <p>TELEPHONE: (202) 408-4000</p> <p>TELEFAX: (202) 408-4400</p> <p>INFORMATION FOR SEQ ID NO: 21:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 569 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: peptide</p> <p>NAME/KEY: Protein</p>			

```

; LOCATION: 1..569
; OTHER INFORMATION: /note= "URE B - FIGURE 3."
; US-08-466-248-21

Query Match          74.8%; Score 2244; DB 2; Length 569;
Best Local Similarity 72.5%; Pred No. 1.8e-39;
Matches 41; Conservative 71; Mismatches 85; Indels 0; Gaps 0;

QY   2   KMKKQEYVNTGYPTKGDKVLGDTDLWAEVEHDYTTYGEELKFAGAKTIRREGMQSNSPD 61
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    3   KISRKEYSVMYGPTTGDRVLGDTLDLILEVEDCTTYGEEKFGGGKTIRDGMSQTSPS 62

QY   62   ENTLLDLVITNAMLIDYTIYKADIGIKNGKIHGIGKAGNKMDQGVSPHVMVGVTGAALA 121
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    63   SYEELDLVLTNALIVDYTYGIYKADIGIKDGTAGIGKAGNKMDQGVNNLCVGPATEALA 122

QY   122   GEGMITTAGDIGDSHTFPLSPQQPPTALANGVTMTMFGGSGTGPVDGTNATTTPGKNLHRM 181
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    123   ARLGLVTAGIGDIHIHFISQQIPTAFASGVTTMGGGTGPADGTNATTTPGRANLKSM 182

QY   182   LRAAEEYSNMNVFLGKGNSSKKQLVPEQVEAGAIGFKLHEDWGVTTPSAIDHCLSVADEYD 241
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    183   LRAAEEYAMNLGFALKGNVSYPESLRQIERAGALGFKIHEWDGSTPAAIHHCLNAVDEYD 242

QY   242   VQVCIIHTDVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVIWAGELNILPSSTTP 301
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    243   VQVAIHDTLTNEACGCVBDTLLEATAGRTIHTFTHTGAGGSHAPDVIKWAGEFNILPASTNP 302

QY   302   TIPYTTINTVAEHLDMLTCHHLDDRIREDFQSRSRIPGSIAAEDVLHDMGVIAWTSSD 361
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    303   TIPFTKNTEAEHMDMLVCHHLDSIKIEDVQFADSRIRPOTIAABDQLHDMGIFSISSD 362

QY   362   SQAMGRAGEVIPRTWQTADNKKBEFGKLDPDGKDNDFRIKRYISKYITINPALTHGVSEY 421
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    363   SQAMGRGVEVITRTWQTADNKKBEFGKLKEEKGDNDNFRIKRYISKYITINPGIAHSIDY 422

QY   422   IGSVEBKIALDLVVNPNAPFGVKPKVIVKGMVVFSEMGSDSNASVPPTQPYYVTEMFGHH 481
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    423   VGSVEVGKYADLVLSPAFFGIKPNNMIKGGFIALSQMGDANASIPTPPQYYYREMFGHH 482

QY   482   GKAKFDTSITFVSKVAVENGVEKGLGERQVLPVKNCRNITKKDKFNKDXTAKITVDPKT 541
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    483   GKNKFDFTNITFVSQAAYKAGIKEELGLDRAAPPVKNCRNITKKDLKENDVTAHIDVNPET 542

QY   542   FEYFVDGKLGCTSKPTSQVPLAQRYTF 568
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    543   YKVVDGKEVTSKADELSLAQLYNLF 569

RESULT 12
US-09-338-920B-10
; Sequence 10, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 569
; TYPE: PRF
; ORGANISM: Helicobacter felis
US-09-338-920B-10

Query Match          74.8%; Score 2244; DB 2; Length 569;
Best Local Similarity 72.5%; Pred. No. 1.8e-39;

```

Matches	411;	Conservative	71;	Mismatches	85;	Indels	0;	Gaps	0;
QY	2	KMKQBYVNTY	GPTGDKVRLGDTDLWAEVEHDYTTYTGELKFGAGKTIREGMSQNSPD	61					
DB	3	KISRKEYVSMYGPTT	GDRVRLGDTDLILEVEHDCTTYGEEIKFGGKTIIRGDMQSTQNSPS	62					
QY	62	ENTLDLVITNAMI	IDYTGIIYKADIGIKNGKIHGIKAGKNQMDQGVSPHMVVGVTGTEALA	121					
DB	63	SYELDLVLNALT	IVDYTGIIYKADIGIKDGKIAGIKAGKNQMDQGVYDNNLCVGPATEALA	122					
QY	122	GEGMITTAGGID	SHTFPLSPQPPALANGVTMTFGGTPVDVGTNATTITPKKNLHRM	181					
DB	123	AEGLITVAGID	THIHFIISQQIPTAFAGVPTMIGGTGTPADGTNATTITPGRANLKSM	182					
QY	182	LRAAEYSNMV	GLKGNSSSKQLVQEVQEAAGAIGFKLHEDMGTTTSAIDHCLLSVADEYD	241					
DB	183	LRAAEYANML	GFLAKGNVSYBPSLRDQIEAGAIGFKIHEDMGSTPAAIHHLNVADEYD	242					
QY	242	VQVCIHTD	TWNEAGYDDTILNAMGRAIHAYHIEGAGGHSPPVITWAGELNLPSSSTTP	301					
DB	243	VQVAIHTDL	TNEAGCYEDTLEATAGRTIHTFTHTEGAGGGHAPDVIKMGAEFNLPASTNP	302					
QY	302	TIPVTINTVAE	HLMDLMTCHHLDKRREDLPQSOSRIRPGSIAAEVDLHDMGVIAMTSSD	361					
DB	303	TIPFTKNTAE	HMDLMTCHHLDSIKEDVQFADSRIRPOTIAAEQLHDMGIFSTSSD	362					
QY	362	SOAMGRAGE	VIPTWTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY	421					
DB	363	SOAMGRVGE	VIPTWTADKNKKEFGKLEKGDNDNFRIKRYISKYITINPGIAHGSY	422					
QY	422	IGSVBEGKI	ADLVVWNPAPFGVQPKVIKGMVVFSEMGDSNASVPTPQPVVYREMFHH	481					
DB	423	VGSVEVGKY	ADLVWSPAPFGIKPNMIKGGFTALSQMGDANASIPTQPVVYREMFHH	482					
QY	482	GKAKFTSIT	FFVSKVAYENGVEKLGLEQVLVVKNCNITTKDKPFNDKTXKITVDPKT	541					
DB	483	GKNKFDNT	ITFFVSOAYKAGIKELGLDRAAPPVVKNCNITTKDKLPNDVTAHIDVNPET	542					
QY	542	FEYFVVDG	KLCCTSKPTSQVPLAQRYTTFP 568						
DB	543	YKVKVDG	KVTSKADELSLAQLYNLF 569						
RESULT 13									
US-09-338-920B-8									
; Sequence 8, Application US/09338920B									
; Patent No. 6709851									
; GENERAL INFORMATION:									
; APPLICANT: Soman, Gopalan									
; APPLICANT: Thomas, William D., Jr.									
; APPLICANT: Monath, Thomas P.									
; TITLE OF INVENTION: Stabilization of Helicobacter Urease									
; FILE REFERENCE: 06132/023002									
; CURRENT APPLICATION NUMBER: US/09/338,920B									
; CURRENT FILING DATE: 2002-12-23									
; PRIOR APPLICATION NUMBER: US 08/928,081									
; PRIOR FILING DATE: 1997-09-12									
; NUMBER OF SEQ ID NOS: 14									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 8									
; LENGTH: 559									
; TYPE: PRT									
; ORGANISM: Helicobacter pylori									
US-09-338-920B-8									
Query Match 74.6%; Score 2237; DB 2; Length 559;									
Best Local Similarity 73.2%; Pred. No. 2.5e-39;									
Matches 406; Conservative 72; Mismatches 76; Indels 1; Gaps 1;									
QY	2	KMKQBYVNTY	GPTGDKVRLGDTDLWAEVEHDYTTYTGELKFGAGKTIREGMSQNSPD	61					
DB	3	KISRKEYVSMYGPTT	GDRVRLGDTDLILEVEHDYTTYTGELKFGGKTIIRGDMQSTQNSPS	62					


```

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26887

Query Match      61.2%; Score 1834; DB 2; Length 566;
Best Local Similarity 59.1%; Pred. No. 6.6e-31;
Matches 337; Conservative 87; Mismatches 140; Indels 6; Gaps 3;

Qy 1 MKMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIREGMGQSNP 60
Db 1 MKISQAYADMFGPTVGRVRLADTLWIEVERDFTVYGEVVKFGGKVI RDGMGSQLG 60

Qy 61 DENTLDLIVITNAMIIDYTIKADIGIKNGKIHGKAGNKMQDGVSPHVMVGVGTEAL 120
Db 61 AAQVVDVTITNALLDHWGVVVKADVGLKDGRIQAIKAGNPDIOQGV--NIAIGAGTEVI 118

Qy 121 AGEGMVITAGGIDSHTHFLSPQFPPTALANGVYTMFGGGTGPVDGTNATTITPGKWNLHR 180
Db 119 AGEGMILTAGGIDTHIHFCPCQIIEALMSGVTTMIGGGTGPAAAGTNATTCTSGPWHWR 178

Qy 181 MLRAAEYSNMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
Db 179 MLQAADAPPNNIGFTGKNASLPLPLEBQVLAGAIGLKHEDWGSTPAAIDNCLVAAERH 238

Qy 241 DVQVCIHDTVNEAGYVDDTLNANNGRAIHAYHIEGAGGHSPPDVTWAGEINILPSSTT 300
Db 239 DIQVAIHDTILNESGFVETTLGAFKGRTIHTYHTGAGGHPADIIKACGFANVLPSSTN 298

Qy 301 PTIPVTINTVAEHLDMLTCHLDKRIREDIQFSQSRIRPGSIAAEDVLHDMGVIAMTSS 360
Db 299 PTRPPTRTIDDEHLDMLVCHHLDPAEDVAFAESRIRRETIAAEDILHDLGAFSMISS 358

Qy 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPBGDKDNDNFRIKRYISKYTIINPALTHGVSE 420
Db 359 DSQAMGRGEVITRTWQTADKNKQGRGLDGDGARNDFRARRYIAKYTIINPAITHGISH 418

Qy 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYREMFCH 480
Db 419 EVGSVEAGKWADLVLRPAFFGVKPSLILKGAIAAASLMGDIINGSIPTPQPVHYRPMFAS 478

Qy 481 HGKAKFDTISITFVSKVAYENGKELGLERQVLVYKNCNITKKDFKENDKTAKITVDPK 540
Db 479 YAGSRHATSLTFVSOAFAAGVPOQLGRKAIGVVSVCGRGVQKTDLIHNGYLPTEIYDAQ 538

Qy 541 TFEVFDGKL--CTSKPTSQVPLAORYTFF 568
Db 539 NYQVRADGQLLWC--EPADVLPMAORYFLF 566

```

Search completed: August 9, 2006, 21:44:42
Job time : 53 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:44:02 ; Search time 185 Seconds

(without alignments)
1422.195 Million cell updates/sec

Title: US-09-904-994B-3

Perfect score: 2999

Sequence: 1 MMKKQEVNTYGTTKGDKV.....KLCTSKPTSQVPLAQRATFF 568

Scoring table: BLOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main.*

1: /EMC_Celerra_SID33/prodata/2/pubpaa/US07_PUBCOMB.pap.*
2: /EMC_Celerra_SID33/prodata/2/pubpaa/US08_PUBCOMB.pap.*
3: /EMC_Celerra_SID33/prodata/2/pubpaa/US09_PUBCOMB.pap.*
4: /EMC_Celerra_SID33/prodata/2/pubpaa/US10A_PUBCOMB.pap.*
5: /EMC_Celerra_SID33/prodata/2/pubpaa/US10B_PUBCOMB.pap.*
6: /EMC_Celerra_SID33/prodata/2/pubpaa/US11_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2999	100.0	568	3	US-09-904-994B-3
2	2973	99.1	568	3	US-09-904-994B-15
3	2960	98.7	568	3	US-09-904-994B-6
4	2953	98.5	568	3	US-09-904-994B-12
5	2591	86.4	496	3	US-09-904-994B-9
6	2326	77.6	569	4	US-10-639-273-5
7	2304	76.8	568	4	US-10-639-273-38
8	2297	76.6	569	3	US-09-995-913A-252
9	2297	76.6	569	4	US-10-282-122A-58611
10	2297	76.6	569	4	US-10-335-977-8611
11	2297	76.6	569	4	US-10-639-273-40
12	2244	74.8	569	4	US-10-639-273-39
13	2239	74.7	750	3	US-09-402-100-2
14	2109	70.3	569	4	US-10-639-273-41
15	1927	64.3	568	4	US-10-282-122A-47638
16	1921	64.1	566	4	US-10-282-122A-69286
17	1913	63.8	568	4	US-10-282-122A-49766
18	1902	63.4	568	4	US-10-282-122A-50238
19	1843	61.5	572	4	US-10-282-122A-58210
20	1838	61.3	566	4	US-10-282-122A-66743
21	1834	61.2	567	4	US-10-282-122A-60007
22	1831	61.1	779	4	US-10-282-122A-69134
23	1826	60.9	566	4	US-10-282-122A-44964
24	1824	60.8	837	4	US-10-424-599-254635
25	1818	60.6	568	4	US-10-282-122A-56823
26	1818	60.6	568	4	US-10-282-122A-56844
27	1810	60.4	567	4	US-10-282-122A-68252

28	1804	60.2	571	5	US-10-793-626-118	Sequence 118, App
29	1804	60.2	811	4	US-10-282-122A-70574	Sequence 70574, A
30	1803	60.1	573	4	US-10-724-972A-5526	Sequence 5526, Ap
31	1797	59.9	571	4	US-10-282-122A-44443	Sequence 44443, A
32	1795	59.9	572	4	US-10-282-122A-78027	Sequence 78027, A
33	1770	59.0	840	4	US-10-621-833-7	Sequence 7, Appli
34	1770	59.0	840	5	US-10-731-877-1	Sequence 1, Appli
35	1770	59.0	840	6	US-11-046-271-1	Sequence 124057, A
36	1769	59.0	843	6	US-10-437-963-124057	Sequence 3599, Ap
37	1728	57.6	570	3	US-09-738-626-3599	Sequence 2, Appli
38	1726	57.6	837	4	US-10-418-962-2	Sequence 76849, A
39	1721	57.4	598	4	US-10-282-122A-76849	Sequence 14, Appl
40	1720	57.4	570	5	US-10-454-437-14	Sequence 51072, A
41	1713	57.1	571	4	US-10-282-122A-51072	Sequence 14635, A
42	1706	56.9	573	4	US-10-156-761-14635	Sequence 16, Appl
43	1631	54.4	534	5	US-10-454-437-16	Sequence 62532, A
44	1624	54.2	577	4	US-10-282-122A-62532	Sequence 64617, A
45	1624	54.2	577	4	US-10-282-122A-64617	

ALIGNMENTS

RESULT 1

US-09-904-994B-3

; Sequence 3, Application US/09904994B

; Publication No. US20040005325A1

; GENERAL INFORMATION:

; APPLICANT: AKZO No. US20040005325A1e1 N.V.

; TITLE OF INVENTION: Helicobacter vaccine

; FILE REFERENCE: Depoedesequences

; CURRENT APPLICATION NUMBER: US/09/904,994B

; CURRENT FILING DATE: 2001-07-13

; NUMBER OF SEQ ID NOS: 21

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 568

; TYPE: PRT

; ORGANISM: Helicobacter felis

US-09-904-994B-3

Query Match 100.0%; Score 2999; DB 3; Length 568;

Best Local Similarity 100.0%; Pred. No. 2.3e-53;

Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MMKKQEVNTYGTTKGDKVRLGDTDLWAEVHDYTTTGEELKFGAGKTIREGMGQSNP	60
Db	1	MMKKQEVNTYGTTKGDKVRLGDTDLWAEVHDYTTTGEELKFGAGKTIREGMGQSNP	60
Qy	61	DENTLDLVITNAMIIDYTGIIYKADIGIKNGKIHGIGKAGNKMDDGVSPhmVVGVTAL	120
Db	61	DENTLDLVITNAMIIDYTGIIYKADIGIKNGKIHGIGKAGNKMDDGVSPhmVVGVTAL	120
Qy	121	AGEGMIITAGIDSHTHPLSQPPTALANGVTWTFGGTGVDTGNATITPGKNLHR	180
Db	121	AGEGMIITAGIDSHTHPLSQPPTALANGVTWTFGGTGVDTGNATITPGKNLHR	180
Qy	181	MLRAEYSNMVGLFGKNSSSKKQLVEQVAGAGIKLHEDWGTTPSAIDHCLSVADY	240
Db	181	MLRAEYSNMVGLFGKNSSSKKQLVEQVAGAGIKLHEDWGTTPSAIDHCLSVADY	240
Qy	241	DVQVCIHDTVNEAGYVDDTLNMGRAIHAYHIEGAGGGHSPDVTWAGELNLPSSYT	300
Db	241	DVQVCIHDTVNEAGYVDDTLNMGRAIHAYHIEGAGGGHSPDVTWAGELNLPSSYT	300
Qy	301	PTIPTYTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEVLHDMGVIAVTSS	360
Db	301	PTIPTYTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEVLHDMGVIAVTSS	360
Qy	361	DSQAMGAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTGHVSE	420
Db	361	DSQAMGAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTGHVSE	420

Qy 421 YIGSVEGKIADLVVWNPAPFGVGPVKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Db 421 YIGSVEGKIADLVVWNPAPFGVGPVKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Qy 481 HGKAKFDTSTIFVSKVAVENGKGLGLERQVLPVKNCRNITKKDFKNDTKAKITVDPK 540
Db 481 HGKAKFDTSTIFVSKVAVENGKGLGLERQVLPVKNCRNITKKDFKNDTKAKITVDPK 540
Qy 541 TFEVFDGKCLCTSKPTSQVPLAQRVTF 568
Db 541 TFEVFDGKCLCTSKPTSQVPLAQRVTF 568

RESULT 2
US-09-904-994B-15
; Sequence 15, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-15

Query Match 99.1%; Score 2973; DB 3; Length 568;
Best Local Similarity 98.9%; Pred. No. 7.9e-53;
Matches 562; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MMKKQEVNTYGTGDKVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMQSNP 60
Db 1 MMKKQEVNTYGTGDKVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMQSNP 60
Qy 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120
Db 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120
Qy 121 AGEGLIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDTGNATTITPGKWNLHR 180
Db 121 AGEGLIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDTGNATTITPGKWNLHR 180
Qy 181 MLRAAEYSMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
Db 181 MLRAAEYSMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
Qy 241 DVQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGELNILPSSTT 300
Db 241 DVQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGELNILPSSTT 300
Qy 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
Db 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
Qy 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
Db 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
Qy 421 YIGSVEGKIADLVVWNPAPFGVGPVKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Db 421 YIGSVEGKIADLVVWNPAPFGVGPVKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Qy 481 HGKAKFDTSTIFVSKVAVENGKGLGLERQVLPVKNCRNITKKDFKNDTKAKITVDPK 540
Db 481 HGKAKFDTSTIFVSKVAVENGKGLGLERQVLPVKNCRNITKKDFKNDTKAKITVDPK 540
Qy 541 TFEVFDGKCLCTSKPTSQVPLAQRVTF 568
Db 541 TFEVFDGKCLCTSKPTSQVPLAQRVTF 568

RESULT 4
US-09-904-994B-12
; Sequence 12, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine

Db 541 TFEVFDGKCLCTSKPTSEVPLAQRVTF 568

RESULT 3
US-09-904-994B-6
; Sequence 6, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-6

Query Match 98.7%; Score 2960; DB 3; Length 568;
Best Local Similarity 98.6%; Pred. No. 1.5e-52;
Matches 560; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MMKKQEVNTYGTGDKVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMQSNP 60
Db 1 MMKKQEVNTYGTGDKVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMQSNP 60
Qy 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120
Db 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120
Qy 121 AGEGLIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDTGNATTITPGKWNLHR 180
Db 121 AGEGLIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDTGNATTITPGKWNLHR 180
Qy 181 MLRAAEYSMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
Db 181 MLRAAEYSMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
Qy 241 DVQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGELNILPSSTT 300
Db 241 DVQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGELNILPSSTT 300
Qy 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
Db 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
Qy 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
Db 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
Qy 421 YIGSVEGKIADLVVWNPAPFGVGPVKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Db 421 YIGSVEGKIADLVVWNPAPFGVGPVKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Qy 481 HGKAKFDTSTIFVSKVAVENGKGLGLERQVLPVKNCRNITKKDFKNDTKAKITVDPK 540
Db 481 HGKAKFDTSTIFVSKVAVENGKGLGLERQVLPVKNCRNITKKDFKNDTKAKITVDPK 540
Qy 541 TFEVFDGKCLCTSKPTSQVPLAQRVTF 568
Db 541 TFEVFDGKCLCTSKPASEVPLAQRVTF 568

RESULT 4
US-09-904-994B-12
; Sequence 12, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine

FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 568
; TYPE: PRN
; ORGANISM: Helicobacter felis
; US-09-904-994B-12

Query Match 98.5%; Score 2953; DB 3; Length 568;
Best Local Similarity 98.2%; Pred. No. 2e-52; Indels 0; Gaps 0;
Matches 558; Conservative 5; Mismatches 5;
QY 1 MKMKQEVYNTYPTKDKVRLGDTDLWAEVEHDYTYGELKFGAGKTIREGMGQNSP 60
DB 1 MKMKQEVYNTYPTKDKVRLGDTDLWAEVEHDYTYGELKFGAGKTIREGMGQNSP 60
QY 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEAL 120
DB 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEAL 120
QY 121 AGSGMIITAGGIDSHTHFLSPQPPPTALANGVTTMFGGGTGPVDGTNATTTTPCKNLHR 180
DB 121 AGSGMIITAGGIDSHTHFLSPQPPPTALANGVTTMFGGGTGPVDGTNATTTTPCKNLHR 180
QY 181 MLRAAEYSNMVGFGLKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
DB 181 MLRAAEYSNMVGFGLKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
QY 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTMAGELNLPSTT 300
DB 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTMAGELNLPSTT 300
QY 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
DB 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
QY 361 DSQAMGRAGEVTPRTWQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
DB 361 DSQAMGRAGEVTPRTWQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
QY 421 YIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYVYREMFH 480
DB 421 YIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYVYREMFH 480
QY 481 HGKAKFDTSTITFVSQVPLAQRVYTF 568
DB 481 TFEVFDGKLTCKSPASEVPLAQRVYTF 568

RESULT 5

US-09-904-994B-9
; Sequence 9, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 496
; TYPE: PRN
; ORGANISM: Helicobacter felis
; US-09-904-994B-9

Query Match 86.4%; Score 2591; DB 3; Length 496;
Best Local Similarity 99.2%; Pred. No. 4.2e-45; Indels 1; Gaps 1;
Matches 491; Conservative 1; Mismatches 2;
QY 1 MKMKQEVYNTYPTKDKVRLGDTDLWAEVEHDYTYGELKFGAGKTIREGMGQNSP 60
DB 1 MKMKQEVYNTYPTKDKVRLGDTDLWAEVEHDYTYGELKFGAGKTIREGMGQNSP 60
QY 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEAL 120
DB 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEAL 120
QY 121 AGSGMIITAGGIDSHTHFLSPQPPPTALANGVTTMFGGGTGPVDGTNATTTTPCKNLHR 180
DB 121 AGSGMIITAGGIDSHTHFLSPQPPPTALANGVTTMFGGGTGPVDGTNATTTTPCKNLHR 180
QY 181 MLRAAEYSNMVGFGLKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
DB 181 MLRAAEYSNMVGFGLKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
QY 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTMAGELNLPSTT 300
DB 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTMAGELNLPSTT 300
QY 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
DB 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
QY 361 DSQAMGRAGEVTPRTWQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
DB 361 DSQAMGRAGEVTPRTWQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
QY 421 YIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYVYREMFH 480
DB 421 YIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYVYREMFH 480
QY 481 HGKAKFDTSTITF-VS 494
DB 481 HGKAKFDTSTITFRVS 495

RESULT 6

US-10-639-273-5
; Sequence 5, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 569
; TYPE: PRN
; ORGANISM: Helicobacter bizzozeronii
; US-10-639-273-5

Query Match 77.6%; Score 2326; DB 4; Length 569;
Best Local Similarity 74.1%; Pred. No. 1.5e-39;
Matches 420; Conservative 75; Mismatches 72; Indels 0; Gaps 0;
QY 2 MKMKQEVYNTYPTKDKVRLGDTDLWAEVEHDYTYGELKFGAGKTIREGMGQNSP 61
DB 3 KISRKEYVSMYPTGTDKVRGLDGLILEVEHDCYTYGSEIKRFGGKTIKRDGMAQTNSP 62

Qy	62	ENTLDLVITNAMIIDYTGICYKADIGIRKNGKIHGIGKAGNKMQDQGVSHMVVGVGTEALA	121
Db	63	SHELDLVITNALIVDYTGICYKADIGIRKNGKIHGIGKAGNKMQDQGVNCLVGPATEALA	122
Qy	122	GEGLMITAGGIDSHTFLSPPOFPTALANGVTTMFGGCTGPDGCTNATITPGKNLHRM	181
Db	123	ABGLIVTAGGIDTHIHFIISPOOIPTAFASGITMIFGGGTGPADGVNATITTPGRWNLKTM	182
Qy	182	LRAABEYSNMVFLGKGNSSKKQIVGEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADEYD	241
Db	183	LRASEEVAMNLGYLGKGNVSEPSLVDQLEAGAIGFKLHEDWGTTPAAYHCLNVADKYD	242
Qy	242	VQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSFDDVTMAGELNLPSTTP	301
Db	243	VQVAIHDTLINEAGCVEDTLQAIAGRTIHTPHTEGAGGCHAPDVIKMSGFNILPASTNP	302
Qy	302	TIPYTIINTVAEHLDMLTCHHLDKRIREDLOFSQSRIRPGSIAABDVLDHMGVIAMTSSD	361
Db	303	TIPFTVNTAEHMDMLVCHHLDKNIKEDVOFADSRIRPQTIAABDKLHDMGIFITSDD	362
Qy	362	SQAMGRAGEVIPRTWOTADNKKRFGKLPEDGKONDNFRIKRYISKYITNPALTGVSEY	421
Db	363	SQAMGRVGEVITRTWQADNKKRFGRLPEBKONDNFRIKRYISKYITNPAIAHIGISEY	422
Qy	422	IGSVEEGKIADLVVWNPFAFFGVPKPIVITKGVVVPSEMGDSNASVPTQPVYVRMFPGH	481
Db	423	VGSVEVGKFADLVLSPAFFGIGKPNWIIKGFIALSQMGDANASITPQPVYVRMFPGH	482
Qy	482	GKAKFDTSITFVSKVAYENGVEKLGLEBROVLVPRKNCRNITKKOPKFNDDTKAKITVDPKT	541
Db	483	GKAKFDNITFVSVQVADNGIKEELGLGRVVLVPRKNCRNITKKOLKFNDDVTAHIEVNPET	542
Qy	542	FEVPVDGLCTSKPSTQVPLAQRITFP	568
Db	543	YKVKVDGKEVTSKAADKISLAQLYNLF	569

```

RESULT 7
US-10-639-273-38
; Sequence 38, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERO
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT OF
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639, 273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404, 337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 568
; TYPE: PRN
; ORGANISM: Helicobacter heilmannii
US-10-639-273-38

```

[illegible][illegible]

RESULT 8

```

US-09-895-913A-252
; Sequence 252, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Omen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1e1 H
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-252

```

	Query Match	76.6%	Score 2297;	DB 3;	Length 569;
	Best Local Similarity	73.4%;	Pred. No. 5.7e-39;		
	Matches 416; Conservative	76;	Mismatches 75;	Indels 0;	Gaps 0
Qy	2	KMKKQEVYNTVPTGDKVRIGDLDLAEVEHDYTYTGEEELKFAGAKTIRMGMSQNSPD	61		
Dd	3	KISRKQVSMYGPTGDKVRIGDLDLAEVEHDYTYTGEEELFGGGKTLREGMSQSNNPS	62		
Qy	62	ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGICKAGNKDMQGVSPHVVGVGEALA	121		
Dd	63	KEELDILITNALIVDYTYGIYKADIGIKDGKIAGIKGKGNDMQGVKNLNVSGVPATEALA	122		
Qy	122	GEGMIITAGGISDTHFLSPQQFTPALANGVTWTFPGGGTGPVDGTNATTTIPGKNKLHRM	181		
Dd	123	GEGLIVTAGGDITHIHFTSPOOIFPAASGVYTMMIGGCTGPADGTNATTTIPGRNLKWM	182		

```
QY 182 LRAAEYSNMVGLKGNSSSKKQVLEAGATGPKLHEDWGTTPSAIDHCLSVADKYD 241
D 183 LRAAEYSNMVGLKGNSSSKKQVLEAGATGPKLHEDWGTTPSAIDHCLSVADKYD 242
QY 242 VQVCIHDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGCHGSPDVTWAGELNIPSSSTP 301
D 243 VQVAIHDTLNEAGCVEDTMAAIAAGRTWMTFTEGAGGCHGAPDIIKVAGEHNILPASTNP 302
QY 302 TIPTVINTVAEHLMLMTCCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVAMTSSD 361
D 303 TIPTVNTAEAHMDMLVCHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSITSSD 362
QY 362 SOAMGRAGEVITPRTWQADKKNKEFGKLPEDCKNDNFRIRKRYLSKYTINPAIAHIGISEY 421
D 363 SOAMGRVGEVITRTWQADKKNKEFGKLPEDCKNDNFRIRKRYLSKYTINPAIAHIGISEY 422
QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVSEMGDSNASVPTPQVYVREMFHGH 481
D 423 VGSVEVGKADLVVWNPAPFGVKPKIVIKGGMVSEMGDSNASVPTPQVYVREMFHGH 482
QY 482 GKAKFDTSTITFVSKVAYENGVEKGLERQVLPVKNCNCRNITKKDFKFNOKTAKITVDPKT 541
D 483 GKAKYDANITFVSOAAYDKGKEELGLERQVLPVKNCNCRNITKKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLCITSKPTSQVPLAQRYTFF 568
D 543 YHVFVDGKEVTSKPANKVSLAQFSIF 569

RESULT 9
US-10-282-122A-58611
; Sequence 58611, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zysek, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58611
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-10-282-122A-58611

Query Match 76.6%; Score 2297; DB 4; Length 569;
Best Local Similarity 73.4%; Pred. No. 5.7e-39;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

QY 2 KMKKQBYVNTYPTGDKVRLGDTDLWAEVEHDYTYTGELKEFGAGKTIREGMSQNSPD 61
D 3 KISRKEYVSMYPTGDKVRLGDTDLIAEVEHDYTYTGELKEFGGKTLREGMSQNSNPS 62
QY 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGKAGNKMOMDGVSPHVVGVGTBALA 121
D 63 KEELDLIITNALIVDYTYGIYKADIGIKDKGIAGIKGKNKMODGVKNLNVGPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGTNATTITPKWNLHRM 181
D 123 GEGLIIVTAGGIDTHIFISPOQIPTAFASGVTTMIGGGTGPADGTNATTITPGRRLKWM 182
QY 182 LRAAEYSNMVGLKGNSSSKKQVLEAGATGPKLHEDWGTTPSAIDHCLSVADKYD 241
D 183 LRAAEYSNMVGLKGNSSSKKQVLEAGATGPKLHEDWGTTPSAIDHCLSVADKYD 242
QY 242 VQVCIHDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGCHGSPDVTWAGELNIPSSSTP 301
D 243 VQVAIHDTLNEAGCVEDTMAAIAAGRTWMTFTEGAGGCHGAPDIIKVAGEHNILPASTNP 302
QY 302 TIPTVINTVAEHLMLMTCCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVAMTSSD 361
D 303 TIPTVNTAEAHMDMLVCHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSITSSD 362
QY 362 SOAMGRAGEVITPRTWQADKKNKEFGKLPEDCKNDNFRIRKRYLSKYTINPAIAHIGISEY 421
D 363 SOAMGRVGEVITRTWQADKKNKEFGKLPEDCKNDNFRIRKRYLSKYTINPAIAHIGISEY 422
QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVSEMGDSNASVPTPQVYVREMFHGH 481
D 423 VGSVEVGKADLVVWNPAPFGVKPKIVIKGGMVSEMGDSNASVPTPQVYVREMFHGH 482
QY 482 GKAKFDTSTITFVSKVAYENGVEKGLERQVLPVKNCNCRNITKKDFKFNOKTAKITVDPKT 541
D 483 GKAKYDANITFVSOAAYDKGKEELGLERQVLPVKNCNCRNITKKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLCITSKPTSQVPLAQRYTFF 568
D 543 YHVFVDGKEVTSKPANKVSLAQFSIF 569

RESULT 10
US-10-335-977-8611
; Sequence 8611, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
```

```
;; SOFTWARE: UNIX
;; CURRENT APPLICATION DATA:
;;   APPLICATION NUMBER: US/10/335,977
;;   FILING DATE: 30-Dec-2002
;; PRIOR APPLICATION DATA:
;;   APPLICATION NUMBER: 08/993,002
;;   FILING DATE: 17-DEC-1997
;; ATTORNEY/AGENT INFORMATION:
;;   NAME: Mandragouras, Amy E.
;;   REGISTRATION NUMBER: 36,207
;;   REFERENCE/DOCKET NUMBER: GTN-018
;; TELECOMMUNICATION INFORMATION:
;;   TELEPHONE: (617)227-7400
;;   TELEFAX: (617)742-4214
;; INFORMATION FOR SEQ ID NO: 8611:
;;   SEQUENCE CHARACTERISTICS:
;;     LENGTH: 569 amino acids
;;     TYPE: amino acid
;;     TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: YES
;; ORIGINAL SOURCE:
;;   ORGANISM: Helicobacter pylori
;; FEATURE:
;;   NAME/KEY: misc feature
;;   LOCATION: (B) LOCATION 1...569
;;   SEQUENCE DESCRIPTION: SEQ ID NO: 8611:
US-10-335-977-8611
```

```
Query Match      76.6%; Score 2297; DB 4; Length 569;
Best Local Similarity 73.4%; Pred. No. 5.7e-39;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

QY      2 KMKQKQEVNTYGTGKDVRLGDTDLAEVHDYTYTGELKFGAGKTIREGMSQNSPD 61
DB      3 KISRKEYVSMYPTTGDKVRLGDTDLAEVHDYTYTGELKFGAGKTIREGMSQNSPS 62

QY      62 ENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGKAGNKMQDGVSPHVVGVGTEALA 121
DB      63 KEELDLIITNALIVDVTGIYKADIGIKGKIAGIKGKNKMDQGVKNLSVGPATEALA 122

QY      122 GEGMIITAGGIDSHVFLSPQOPFALANGVTTMTFGGCTGPDVGTNATTITPGKNLHRM 181
DB      123 GEGLIIVTAGGIDTHIFISPOQIPTAFASGVTTMIGGGTGPADGTNATTITPGRENKWM 182

QY      182 LRAAEYSNMVGLGKGNSSKKQVQVEAGAIGFKLHEDMGTTPSAIDHCLSVADVD 241
DB      183 LRAAEYSNMVGLGKGNSSKKQVQVEAGAIGFKLHEDMGTTPSAIDHCLSVADVD 242

QY      242 VQVCITHTDVTNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPOVITWAGELNIPSTTP 301
DB      243 VQVAITHTDVTNEAGVEDTMAAIGRTWHTPHTGAGGGHAPDIIKVAGEHNILPASTNP 302

QY      302 TIPTVTINTVAEHLDMMLTCHLDKRIREDLPQSQRIRPGSIAAEDVLHDMGVIAWTSDD 361
DB      303 TIPTVTNTAEHMDMLVCHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSISSD 362

QY      362 SOAMGRAGEVTPRTWQADKNKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 421
DB      363 SOAMGRVGEVITRTWQADKNKEFGRLKEEKGNDNFRIKRYLSKYITINPAIAHGISEY 422

QY      422 IGSVEBGKIADLVVNPAPFFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVYREMPFGH 481
DB      423 VGSVEVGKADLVLSVPAFFGVKPNMIIKGGFIALSQMGDANASITPTPOPVYVYREMPFAH 482

QY      482 GKAKFDTSITFVSKVAYENGKVLGLERQVLVFNKCRNITKKDKFPNDKTKAKITVDPKT 541
DB      483 GKAKYDANITFVSQAAYDKGIEELGLERQVLVFNKCRNITKKDMQFNDTTAHIEVNPET 542

QY      542 FEVFDGKLTCKTSKPTSOVPLAQRYTFF 568
DB      543 YHVFVDGKEVTSKPANKVSLAQFLSFIF 569
```

```
RESULT 12
US-10-639-273-39
; Sequence 39, Application US/10639273
; Publication No. US20040142343A1
```

```
RESULT 11
US-10-639-273-40
; Sequence 40, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
;   APPLICANT: Chang, Yung-Fu
;   APPLICANT: Simpson, Kenneth W
;   APPLICANT: Zhu, Jiaqian
;   TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
;   TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
;   FILE REFERENCE: 19603/3881
;   CURRENT APPLICATION NUMBER: US/10/639,273
;   CURRENT FILING DATE: 2003-08-12
;   PRIOR APPLICATION NUMBER: 60/404,337
;   PRIOR FILING DATE: 2002-08-16
;   NUMBER OF SEQ ID NOS: 41
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 40
;   LENGTH: 569
;   TYPE: PRN
;   ORGANISM: Helicobacter pylori
US-10-639-273-40
```

```
Query Match      76.6%; Score 2297; DB 4; Length 569;
Best Local Similarity 73.4%; Pred. No. 5.7e-39;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

QY      2 KMKQKQEVNTYGTGKDVRLGDTDLAEVHDYTYTGELKFGAGKTIREGMSQNSPD 61
DB      3 KISRKEYVSMYPTTGDKVRLGDTDLAEVHDYTYTGELKFGAGKTIREGMSQNSPS 62

QY      62 ENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGKAGNKMQDGVSPHVVGVGTEALA 121
DB      63 KEELDLIITNALIVDVTGIYKADIGIKGKIAGIKGKNKMDQGVKNLSVGPATEALA 122

QY      122 GEGMIITAGGIDSHVFLSPQOPFALANGVTTMTFGGCTGPDVGTNATTITPGKNLHRM 181
DB      123 GEGLIIVTAGGIDTHIFISPOQIPTAFASGVTTMIGGGTGPADGTNATTITPGRENKWM 182

QY      182 LRAAEYSNMVGLGKGNSSKKQVQVEAGAIGFKLHEDMGTTPSAIDHCLSVADVD 241
DB      183 LRAAEYSNMVGLGKGNSSKKQVQVEAGAIGFKLHEDMGTTPSAIDHCLSVADVD 242

QY      242 VQVCITHTDVTNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPOVITWAGELNIPSTTP 301
DB      243 VQVAITHTDVTNEAGVEDTMAAIGRTWHTPHTGAGGGHAPDIIKVAGEHNILPASTNP 302

QY      302 TIPTVTINTVAEHLDMMLTCHLDKRIREDLPQSQRIRPGSIAAEDVLHDMGVIAWTSDD 361
DB      303 TIPTVTNTAEHMDMLVCHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSISSD 362

QY      362 SOAMGRAGEVTPRTWQADKNKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 421
DB      363 SOAMGRVGEVITRTWQADKNKEFGRLKEEKGNDNFRIKRYLSKYITINPAIAHGISEY 422

QY      422 IGSVEBGKIADLVVNPAPFFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVYREMPFGH 481
DB      423 VGSVEVGKADLVLSVPAFFGVKPNMIIKGGFIALSQMGDANASITPTPOPVYVYREMPFAH 482

QY      482 GKAKFDTSITFVSKVAYENGKVLGLERQVLVFNKCRNITKKDKFPNDKTKAKITVDPKT 541
DB      483 GKAKYDANITFVSQAAYDKGIEELGLERQVLVFNKCRNITKKDMQFNDTTAHIEVNPET 542

QY      542 FEVFDGKLTCKTSKPTSOVPLAQRYTFF 568
DB      543 YHVFVDGKEVTSKPANKVSLAQFLSFIF 569
```



```

; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; FILE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Helicobacter hepaticus
; US-10-639-273--41

Query Match          70.3%   Score 2109; DB 4; Length 569;
Best Local Similarity 68.1%; Pred No. 4.le-35;
Matches 388; Conservative 83; Mismatches 95; Indels 4; Gaps 4;

QY 1 M M K K Q S Y V N T Y G P T T G K D K V R L G D T D L W A E V E H D Y T T Y G B E L K F G A G K T I R E G M G Q S N S P 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 I K I S R Q V A S M Y G P T T G C K V R L G D T N L F A B I E K D Y T L Y G E E I F G G K T I R D G M A Q S A S T 61

QY 61 D E N T L D L V I T N A M I D Y T G I Y K A D I G I K N G K I H I G I K A G N K M O D Q V S P H M V G V G T E A L 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 Y T N E L D A V I T N A M I D Y T G I Y K A D I G I K G G I K V I G I K A G N P D T Q D S Y N E A M V V G A A T E V I 121

QY 121 A G R G M I I T A G G I D S H T F L S P Q P P T A L A N G V T T M F G S G T G P V D G T N A T T I T P C K W N L H R 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 A G R G Q I I T A G G I D T H I F I S P T I P T A L Y S G V T T M I G G G T G P A A G T N A T T C T P G K W N N H Q 181

QY 181 M L R A A E Y S M N V G F L G K G N S S K Q L V E Q V E A G A I G F K L H E D W C T T P S A I D H C L S V A D E Y 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 M L R A A S Y A M L G F F G K G N S N E E G L E Q I K A G A L G L K V H E D W G S T P A A I N H A L N V A Q Y 241

QY 241 D V Q C I H T D T V N B A G Y V D D T L N A M N G R A I H A Y H I E G A G G H S P D V I T W A G E L N I L P S T T 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 242 D V Q V A I H T D L N B A G C V E D T M K A I D G R T I H T F H T E G A G G H G A P D I I K A A G E P N I L P A S T N 301

QY 301 P T T P Y T I N T V A E H L D M L M T C H I L D K R I R E D L Q S Q S R I R P C S I A A E D V L H D M G V I A M T S S 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 302 P T I P F T K N T A D E H L D M L M V C H I L D K K I K E D V A F A D S R I R P E T I A A E D T L H D M G I F S I T S S 361

QY 361 D S Q A M G R A G E V I P R T W T O T A D K N K E F G K L P E D - G K D N D N F R I K Y I S K Y T I N P A L T H G V S 419
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 362 D S Q A M G R V G E V I P R T W T O T A D K N K E F G A L K E E C G - E N D N F R I K Y I S K Y T I N P A I A N G I S 420

QY 420 E Y I G S V E B G K I A D L V V M N P A F G V K P K I V I K G M V V F S E M G D S N A S V P T P Q P V Y Y R E M F G 479
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 E Y V G S V E V G K F A D L V L M K P S M F G I K P E M I L K N G M I V A A K I G D S N A S I P T P E P V V Y A P M F G 480

QY 480 H H G A K A P D S I T F V S K V A Y E N G V K E K I G L E R Q V L P V N C R N I T K D K F P N D K T A K I T V D P 539
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 481 S Y G K A K Y C A I T F V S K T A Y D C H I K E E L G L E R I L L P V N C R N I T K D M K F N D V I T P I E V N P 540

QY 540 K T E F E V F D G - K L C T S K P T S O V P L A Q R Y T F F 568
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 541 E T Y E V R V N N T K I - T S K P E V K V S L G L Y C L P 569
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-10-282-122A-47638
; Sequence 47638, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Heselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel

```

```

; APPLICANT: Trawick, John
; APPLICANT: Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIORITY APPLICATION NUMBER: 60/191,078
; PRIORITY FILING DATE: 2000-03-21
; PRIORITY APPLICATION NUMBER: 60/206,848
; PRIORITY FILING DATE: 2000-05-23
; PRIORITY APPLICATION NUMBER: 60/207,727
; PRIORITY FILING DATE: 2000-05-26
; PRIORITY APPLICATION NUMBER: 60/230,335
; PRIORITY FILING DATE: 2000-09-06
; PRIORITY APPLICATION NUMBER: 60/230,347
; PRIORITY FILING DATE: 2000-09-09
; PRIORITY APPLICATION NUMBER: 60/242,578
; PRIORITY FILING DATE: 2000-10-23
; PRIORITY APPLICATION NUMBER: 60/253,625
; PRIORITY FILING DATE: 2000-11-27
; PRIORITY APPLICATION NUMBER: 60/257,931
; PRIORITY FILING DATE: 2000-12-22
; PRIORITY APPLICATION NUMBER: 60/267,636
; PRIORITY FILING DATE: 2001-02-09
; PRIORITY APPLICATION NUMBER: 60/269,308
; PRIORITY FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47638
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
; US-10-282-122A-47638

Query Match          64.3%; Score 1927; DB 4; Length 568;
Best Local Similarity 61.3%; Pred. No. 2.2e-31;
Matches 348; Conservative 97; Mismatches 121; Indels 2; Gaps 1;

Qy 1  MWMKKGEYVNTYGTGDKVRLGDTDLWAEVHEHDYTTYTGSEELKFGAGKTIREGMGQSNP 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3  LRLSRRAYAEFMFGTTGDRVRLADTELLIEIRDFTTYGSEVFGGKVIIRDGMGQSRV 62

Qy 61  DENTLDLIVTNAMIIDYTYGYKADIGIKNGKIHGICKAGNKQMDGVSPHWVGVGTEAL 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63  AADVPPTVTINAVILDHGWGVKADIAIKHGRIAAIGKAGNPDIQPGVT--IAIGAATEVI 120

Qy 121  AGSGMIITAGGIDSHTHFTLSPQOFPPTALANGVTTMFGGGTGPVDGTTNATTITPGKWNLHR 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121  AGSLIVTAGGIDTHLHFTSPQOIDEALASGVTTMLGGTGATGATGNATTCTPGPWNER 180

Qy 181  MLRAAEYSMNVGFLGKGNSSKQLVEQVACAGIAGFKLHEDWGTTTPSAIDHCLSVADEY 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181  MLQAADGWPINLGLGKGNASLPQPLVEQIAAAGIAGLKLHEDWGTTTPAADNCLSVADDT 240

Qy 241  DVQVCITHDTVNEAGVYDDTLNANMGRAIHAYHIEGAGGSHSPDIVTMAGELNLPSITT 300
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241  DTQVAIHDTDLNEAGVFVESTVAAFKGRTHYHTGAGGGGHAPDILKVCGEMNVLPSSTN 300

Qy 301  PTPYTIINTVAEHLDMLTCHLDKRIREDLOFSQSRIRPGSIAAEDVLHDMGVAMTSS 360
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301  PTPYTIINTLDEHLDMLMVCCHLDPSIAEDLAFESIRIRRETTAAEDILHDLGNALMSLS 360

Qy 361  DSQAMGRAGEVIRPTWQTADKNKFEKGLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361  DSQAMGRGVEVIRPTWQTAHKMKVQRGALPEDTARNDNFEAKRYVAKYITINPALTHGIAH 420

Qy 421  YIGSVBEGKIADLVWNPAPFGVKPKIVLKGVMVFPSEMGDSNASVPTPOPVYREMPGH 480
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421  EVGSIIPGKWADLVLEPAPFGIKPSMIKGGMIALAQMGDPNASIPTPOVHYREMFAT 480

```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:44:57 ; Search time 34 Seconds
(without alignments)
1124.507 Million cell updates/sec

Title: US-09-904-994b-3

Perfect score: 2999

Sequence: 1 MKMKQEVNTYGTGDKV.....KLCTSKPTSQVPLAQRYYTF 568

Scoring table:

Gapop 4.0 , Gapext 1.0

Searched: 239914 seqs, 67312017 residues

Total number of hits satisfying chosen parameters: 239914

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US09_NEW_PUB_PUB.pap:*
- 2: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US06_NEW_PUB_PUB.pap:*
- 3: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US07_NEW_PUB_PUB.pap:*
- 4: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US08_NEW_PUB_PUB.pap:*
- 5: /EMC_Celerra_SIDS3/prodata/2/pubpaa/PCT_NEW_PUB_PUB.pap:*
- 6: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US10_NEW_PUB_PUB.pap:*
- 7: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US11_NEW_PUB_PUB.pap:*
- 8: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US60_NEW_PUB_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1797	59.9	571	6	US-10-471-571A-3300 Sequence 3300, Ap
2	1794	59.8	653	6	US-10-953-349-33827 Sequence 33827, A
3	1794	59.8	784	6	US-10-953-349-33826 Sequence 33826, A
4	1794	59.8	841	6	US-10-953-349-33825 Sequence 33825, A
5	1769	59.0	843	6	US-10-449-902-47195 Sequence 47195, A
6	1748	58.3	836	7	US-11-292-431-37 Sequence 37, Appl
7	316	10.5	2314	7	US-11-013-711-11 Sequence 11, Appl
8	315	10.5	572	6	US-10-505-928-98 Sequence 98, Appl
9	315	10.5	572	7	US-11-301-554-1815 Sequence 1815, Ap
10	309	10.3	1629	6	US-10-471-571A-3542 Sequence 3542, Ap
11	308	10.3	1006	6	US-10-449-902-41462 Sequence 41462, A
12	308	10.3	1705	7	US-11-304-590-3 Sequence 3, Appl
13	305	10.2	7362	7	US-11-330-403-9773 Sequence 9773, Ap
14	303	10.1	801	6	US-10-471-571A-5160 Sequence 5160, Ap
15	303	10.1	4834	6	US-10-505-928-827 Sequence 827, App
16	302	10.1	636	6	US-10-471-571A-1606 Sequence 1606, Ap
17	301	10.0	1349	6	US-10-471-571A-3352 Sequence 3352, Ap
18	301	10.0	2053	7	US-11-013-711-9 Sequence 9, Appl
19	299	10.0	753	6	US-10-449-902-45171 Sequence 45171, A
20	298	9.9	698	6	US-10-953-349-4543 Sequence 4543, Ap
21	298	9.9	709	6	US-10-953-349-4542 Sequence 4542, Ap
22	298	9.9	725	6	US-10-953-349-4541 Sequence 4541, Ap
23	297	9.9	1992	7	US-11-013-711-3 Sequence 3, Appl
24	297	9.9	1992	7	US-11-013-711-13 Sequence 13, Appl
25	297	9.9	2047	7	US-11-013-711-4 Sequence 4, Appl

26	297	9.9	2047	7	US-11-013-711-7	Sequence 7, Appl
27	295	9.8	951	6	US-10-449-902-41597	Sequence 41597, A
28	295	9.8	1247	7	US-11-330-403-18943	Sequence 18943, A
29	295	9.8	1272	7	US-11-330-403-3001	Sequence 3001, Ap
30	294	9.8	546	7	US-11-330-403-14342	Sequence 14342, A
31	294	9.8	1084	7	US-11-121-154-115	Sequence 115, App
32	293	9.8	928	7	US-11-330-403-10515	Sequence 10515, A
33	293	9.8	2499	7	US-11-070-573-26	Sequence 26, Appl
34	292	9.7	574	7	US-11-330-403-17416	Sequence 17416, A
35	292	9.7	1531	7	US-11-333-747A-44	Sequence 44, Appl
36	291	9.7	539	6	US-10-449-902-50636	Sequence 50636, A
37	291	9.7	574	7	US-11-330-403-9966	Sequence 9966, Ap
38	291	9.7	987	6	US-10-449-902-45475	Sequence 45475, A
39	291	9.7	1167	7	US-11-121-154-1	Sequence 1, Appl
40	290	9.7	563	7	US-11-056-355B-83788	Sequence 83788, A
41	290	9.7	565	7	US-11-056-355B-83787	Sequence 83787, A
42	290	9.7	566	7	US-11-056-355B-83786	Sequence 83786, A
43	290	9.7	2740	7	US-11-330-403-5490	Sequence 5490, Ap
44	290	9.7	3016	7	US-11-330-403-7229	Sequence 7229, Ap
45	290	9.7	22152	6	US-10-544-944-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-10-471-571A-3300
; Sequence 3300, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:

; APPLICANT: CHIRON Spa

; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS

; FILE REFERENCE: P026927W0

; CURRENT APPLICATION NUMBER: US/10/471,571A

; CURRENT FILING DATE: 2003-09-12

; PRIOR APPLICATION NUMBER: GB-0107661.1

; PRIOR FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 5642

; SOFTWARE: SeqWin99, version 1.03

; SEQ ID NO 3300

; LENGTH: 571

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

; FEATURE:

; NAME/KEY: MISC FEATURE

; LOCATION: (1)..(571)

; OTHER INFORMATION: urease alpha subunit

US-10-471-571A-3300

Query Match	59.9%	Score 1797;	DB 6;	Length 571;
Best Local Similarity	57.5%	Pred. No. 1.1e-35;		
Matches	330;	Conservative	97;	Mismatches 134; Indels 13; Gaps 6;
Qy	2	MMKKQEVNTYGTGDKVKRLGDTDLMAVEHDYTYGHELFKFGAGKTTREGMGQSNP-	60	
Db	4	KMTQNTSLYGTGDSIRLGDTNLFQAIEKDYAVYGEATFGGKSRIDGMAQ-N-PR	61	
Qy	61	-----DENTLDIVITNMIIDYTYGKADIGIKNGKHIGIKAGNKMDQGVSPHVVGVG	116	
Db	62	VTRDDVNVADIVISNAVIIIDYKVKADIGIKNGYIFAIGNAGNPDIMDND--IIIGST	119	
Qy	117	TEALAGEMIITAGIDISHTHFLSPQOPPTALANGVTTWFGGGTGPVDCGTNATTITPGKW	176	
Db	120	TDITAAEGKIVTAGIDTHVHFINPEQAEVALESITTHIGGTCASEGSKATVTPGFW	179	
Qy	177	NLHRLRAABEYSMMNVGLFGKNGSSKKQLVQVEAGAIGFKLHEDWGTPPSAIDHCLSV	236	
Db	180	HIHRLAEAGELPINVGFTGKQATNPALTIEQINAGAIGLVKVEDWGATPSALSHALDV	239	
Qy	237	ADEYDVQVCHTDTVNEAGYVDDTLNANNGRAIHAYHLEGAGGHSPPDVTWAGELNILP	296	
Db	240	ADEFDVQIALHALDTLNEAGFMEDTMAAVKDRVLVHYHTEGAGGGHAPDLIKSAAFSNILP	299	

```
QY 297 SSTPTTPTINTVAEHLDMATCHHLDKRIRDLQFSQSRIRPGSIAAEDVLHDMGVIA 356
Db 300 SNTPTLPTTNTVDEHLDVMYMI THLNAALPEDIAFADSRIRKETIAAEDVLQDMGVFS 359
QY 357 MTSSDSQAMGRAGVIRPTWOTADKNKKEFGKLPEDG--KDNDFRIKRYISKYITINPAL 414
Db 360 MISSDSQAMGRVGEVITRTQVAHRMKEORG--PUDGDFFHNDNRIRKRYIAKYITINPAI 417
QY 415 THGVSEYIGSVBEKGIADLVVNPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVY 474
Db 418 THGISVYGSIEPKLADLVLDPIFFGVGPVLPVVGGLINSVANGDANGSIPTSEPMKY 477
QY 475 REMEGHGHKAKFDTSITFVSKVAVENGVEKGLGLERQVLPVKNCRNITTKDFKFNDKTAK 534
Db 478 RKMVGQYGGNLTSTSMTFVSKTAYENGINRALNKRMYRVPVKNIRQLSKADMKNNSATPK 537
QY 535 ITVDPKTFEVPDVGKLCSTKPTSOVPLAQRVTF 568
Db 538 LDVDPQIEVTVGGEKITNSNAETBLPTQRYFLP 571

RESULT 2
US-10-953-349-33827
; Sequence 33827, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33827
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33827

Query Match 59.8%; Score 1794; DB 6; Length 653;
Best Local Similarity 58.9%; Pred. No. 1.8e-35;
Matches 335; Conservative 89; Mismatches 139; Indels 6; Gaps 2;

QY 6 QEYVNTYPTKDGKVRGLDGLDLMWAEVEHDYTYGELKFGAGKTIREGMQO--SNSPDENT 64
Db 85 EKYSIYGTGDKIRLGTNLNLYAEIKDFAFYGDCEIFGGGKVLDRDGMGQASGYPESFC 144
QY 65 LDLVITNAMIIDYTYGIYKADIGIKNGKIHIKAGNKMDDGVSPHMVVGVTGTEALAGEG 124
Db 145 LDTVITNAVVIDYTYGIYKADIGIKGLIIVAIKAGNPDVMDGVHNNMIVGVNTEVIASEG 204
QY 125 MIITAGGIDSHFLSPQOFPALANGVTTFMGCGTGPVDTGNATTITPKGNLHRLRA 184
Db 205 MIVTAGGIDCHVHFCPLQABEAIASGITTLVGGGTGPAHGTCATTCTPAPSQKLMLQS 264
QY 185 ABEYSNMVGFLLKGNSSKKQLVEQVGAIGFKLHEDWGTTPSAIDHCLSVADVDVQV 244
Db 264 TDQPLINMGFTKGNTSKPEGLAEIIKAGAMGLKLFHEDWGTTPSAIDNCLSVADFDIQV 324
QY 245 CIHTDTVNEAGYVDDTLNAMGRAIHAYHIEGAGGHSPPDVTIMAGELNLPSSTTPTIP 304
Db 324 TDQPLINMGFTKGNTSKPEGLAEIIKAGAMGLKLFHEDWGTTPSAIDNCLSVADFDIQV 384
QY 305 YTIINTVAEHLDMATCHHLDKRIRDLQFSQSRIRPGSIAAEDVLHDMGVIAVMTSSDQA 364
Db 385 FTSNTVDEHLDMLVCHHLDKNIPEDVAFESRIRAEITAAEDILHDMGAISIISSDQA 444
QY 365 MGRAGEVIRPTWOTADKNKKEFGKLPEDG-----KDNDFRIKRYISKYITINPALTHGV 419
Db 445 MGRGEVIRTRTWTANKMKVQSGSLPGSGDANAAPDSNLRIRRYIAKYITINPAIVNGFS 504
QY 420 EYIGSVBEKGIADLVVNPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYREMF 479
```

```
Db 505 DFVGSVEVGKADLVLMKPSFFGAKPELVKGAIAWANNGDPNASIPTPEPVMRPMFG 564
QY 480 HHGKAFDTSITFVSKVAYENGVEKGLGLERQVLPVKNCRNITTKDFKFNDKTAKITVDP 539
Db 565 AFGKAGSSNSIAFVSKAAKEAGVATEYRLERKVEAVGRVRLTKLDMKLNALPKIEVDP 624
QY 540 KTFEYVFDGKLCSTKPTSOVPLAQRVTF 568
Db 625 ETYTVDAGDEVLTCPAPTLPLSRNYFLP 653

RESULT 3
US-10-953-349-33826
; Sequence 33826, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33826
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33826

Query Match 59.8%; Score 1794; DB 6; Length 784;
Best Local Similarity 58.9%; Pred. No. 2.7e-35;
Matches 335; Conservative 89; Mismatches 139; Indels 6; Gaps 2;

QY 6 QEYVNTYPTKDGKVRGLDGLDLMWAEVEHDYTYGELKFGAGKTIREGMQO--SNSPDENT 64
Db 216 EKYSIYGTGDKIRLGTNLNLYAEIKDFAFYGDCEIFGGGKVLDRDGMGQASGYPESFC 275
QY 65 LDLVITNAMIIDYTYGIYKADIGIKNGKIHIKAGNKMDDGVSPHMVVGVTGTEALAGEG 124
Db 276 LDTVITNAVVIDYTYGIYKADIGIKGLIIVAIKAGNPDVMDGVHNNMIVGVNTEVIASEG 335
QY 125 MIITAGGIDSHFLSPQOFPALANGVTTFMGCGTGPVDTGNATTITPKGNLHRLRA 184
Db 336 MIVTAGGIDCHVHFCPLQABEAIASGITTLVGGGTGPAHGTCATTCTPAPSQKLMLQS 395
QY 185 ABEYSNMVGFLLKGNSSKKQLVEQVGAIGFKLHEDWGTTPSAIDHCLSVADVDVQV 244
Db 396 TDQPLINMGFTKGNTSKPEGLAEIIKAGAMGLKLFHEDWGTTPSAIDNCLSVADFDIQV 455
QY 245 CIHTDTVNEAGYVDDTLNAMGRAIHAYHIEGAGGHSPPDVTIMAGELNLPSSTTPTIP 304
Db 456 NIHTDTLNEGCVHEHTIAAFKGRAIHTYHSEGAGGHPDIIKCVGKNVLPSSNTPTP 515
QY 305 YTIINTVAEHLDMATCHHLDKRIRDLQFSQSRIRPGSIAAEDVLHDMGVIAVMTSSDQA 364
Db 516 FTSNTVDEHLDMLVCHHLDKNIPEDVAFESRIRAEITAAEDILHDMGAISIISSDQA 575
QY 365 MGRAGEVIRPTWOTADKNKKEFGKLPEDG-----KDNDFRIKRYISKYITINPALTHGV 419
Db 576 MGRGEVIRTRTWTANKMKVQSGSLPGSGDANAAPDSNLRIRRYIAKYITINPAIVNGFS 635
QY 420 EYIGSVBEKGIADLVVNPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYREMF 479
Db 636 DFVGSVEVGKADLVLMKPSFFGAKPELVKGAIAWANNGDPNASIPTPEPVMRPMFG 695
QY 480 HHGKAFDTSITFVSKVAYENGVEKGLGLERQVLPVKNCRNITTKDFKFNDKTAKITVDP 539
Db 696 AFGKAGSSNSIAFVSKAAKEAGVATEYRLERKVEAVGRVRLTKLDMKLNALPKIEVDP 755
QY 540 KTFEYVFDGKLCSTKPTSOVPLAQRVTF 568
```



```
; APPLICANT: XUE, Jiamin
; APPLICANT: YU, Jieh-Juen
; TITLE OF INVENTION: Attenuated Vaccine Useful for
; TITLE OF INVENTION: Immunizations Against Coccidioides spp. Infections
; FILE REFERENCE: 52952200600
; CURRENT APPLICATION NUMBER: US/11/292,431
; CURRENT FILING DATE: 2005-12-02
; PRIOR APPLICATION NUMBER: US 60/633,399
; PRIOR FILING DATE: 2004-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Coccidioides posadasii
US-11-292-431-37

Query Match      58.3%; Score 1748; DB 7; Length 836;
Best Local Similarity 57.2%; Pred. No. 4.2e-34;
Matches 327; Conservative 93; Mismatches 136; Indels 16; Gaps 10;

Qy      3 MKKQEVVNTYPTKGDVRLGDTDLWAEVHDYTYGEEKFGAGKTIREGMGQS---NS 59
Db      272 MTRAYARMFGTTGDDVVGKGTDLWIKVEKDLTYGDECFGGKTIIRDGMGQATGRHS 331

Qy      60 PDENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEA 119
Db      332 VD--VLDTVLVNALIVDMTYGIYKADIGLKGDLGKAGNPDMDGVTPNNIVGSSTDV 389

Qy      120 LAGEMIITAGDISHTHFLSPQPPPTALANGVTMTFGGGTGPVDGVTNATTIPGKNLH 179
Db      390 IACEGKIVTAGIDITHVHFICPPQVEEALASGVITLLGGGTGPTGEGSNATTCTPAP-NQF 448

Qy      180 R-MLRAAEYSNMGFLGKNSKKQLVEQVAGAIKFLHEDWGTTPSAIDHCLSVAD 238
Db      449 KTMQACDHLPLNVLGTGKNDGSLPSLRDQCRAGAGLKVHEDWGATPAVIDTCLQVCD 508

Qy      239 EYDVQVC-IHTDTVNEAGYVDLTNMMNGRAIHAYHIEGAGGHSPPDVTIMAGELNIPLS 297
Db      509 EFDIQ-CLIHDTLINESGFVQCTNAFKRVVHTYHTEGAGGHAPDIISVVEKENVLPS 567

Qy      298 STTPTPTVINTVAHLDMLTCHHLDKRIRDLQFSQSRIRPGSIAAEDVLHDMGVIAM 357
Db      568 STNTPRPTVNTLDEHLDWVCHHLSKDIPEDAFASRISRTIAAEDVLHDTGALSM 627

Qy      358 TSSDSQAMGRAGEVTPRTWQADNKKKFKLPED-GKDNDFRIKRVISKYTTINPALTH 416
Db      628 LSSDSQAMGRGCEVVVVTWNTAHKNMERGRGLKEDEGTDSDNFRVKRYISKYTTINPAIAQ 687

Qy      417 GVSEYIGSEVGKADLVVWNPAPFGVVKPKIVIKGMVVFSEMGDSNASVPTPQPVYRE 476
Db      688 GMAHTIGSEVGKADLVWFPANFGTTPSWLKSMAVSAQMDPNQSGIPTIEPIWRP 747

Qy      477 MFGH-HGKAKFDTSTIFTSKVAYENGVEKGLGLERQVLVFNCRNITKKDPKFNDKTAKI 535
Db      748 MYASLNPR-----SIMFVSQASIKLIGDSYHLKRIEIPVKNCRNISKRDMKFNIMPKN 803

Qy      536 TVDPKTFEYVDGKLTCKTSKPTSQVPLAQRYTF 567
Db      804 RVDPEYVVEADGEBCTEAPVSELPLOTQY-F 834

RESULT 7
US-11-013-711-11
; Sequence 11, Application US/11013711
; Publication No. US20060159709A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; TITLE OF INVENTION: PROTEIN OF MORAXELLA
```

```
; FILE REFERENCE: 1038-921MIS:Jb
; CURRENT APPLICATION NUMBER: US/11/013,711
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/09/361,619
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2314
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-711-11

Query Match      10.5%; Score 316; DB 7; Length 2314;
Best Local Similarity 25.7%; Pred. No. 2.8e+02;
Matches 196; Conservative 83; Mismatches 215; Indels 270; Gaps 90;

Qy      1 MKMKQEVVNTYPTKGD-KV-RLGD-TDLWAEVHDYTYGEEKL-----FGA--GKT 49
Db      587 VKITKDSGINA-----GDQKISNVKDATD-----DTDAVTV-KOLKQVQDADGALQSFS 635

Qy      50 IREGMGQ-----SNSPDENTLDLVIT-----NAMII--DVT-GIYKADIGIK--NG-- 90
Db      636 IRDEKQGEFTISNLYSNGNTPNTPF-TTIFAGENGISINDIAGK--KVKVGIDPINGLT 692

Qy      91 --KIH-GI---GK-----AGN--KDMQDGVSPHM--VVGVG-----TEALAGEGMII 127
Db      693 TPKLVGSDKDGKTKLVIEQVAGSNDTKNIIRGLSPTLPSITNAGGVRTTE---QGNII 748

Qy      128 T-----AGGI-D-SHTHF-LSPQFPPTALANGVTMT-FGGGTGPVDGTNATT--ITPG 174
Db      749 TSEDEKSKAAASIGDILNTGFLNKNNSVGFSTYNTVDF-----IDG-NATTAKVTYD 801

Qy      175 KNLHRLMRAAE-EYSMN-----GFLGKNSSSKKQLVEQVAGAIKFLHEDWGT 225
Db      802 ETN-----QTSKVTVYDVNVDEKTIETLTDGNGKTKIGVK-TTTLTTTNANG-KA-TNFST 853

Qy      226 TPS-AIDHCLSVADYDVQVC-IHT-----DT-----VNEAGYVDD-TL-----NAMNG 266
Db      854 TDNDALVNAKOIAENLNTLAKIEHTTKGTADTALQTFKVKDGTATDEITITVGKDGTONG 913

Qy      267 RAIHAYHIEGAG-----GHSPD-VITM-----AGE---LN-----IL-P----- 296
Db      914 KTVNTLKLKGENGLVATNKGCTVTFGINTQSLKAGDSTTLNKGSLIKNPASNEQIQV 973

Qy      297 -----SST-----TPTIPTYT-----IN-TVAE-HLDMLTCHHL--DK-RIED 330
Db      974 GADGVKFAKVDKNSSTGIDGTSRI--TKDQIGFTGANGSLD--TTKPHLTDKLKVGE- 1028

Qy      331 LQFSQSRIRPGSIAAEDVLHDMGVIAWTSDDSOAMGRA-----GEV---IPRTWQTADKNK 383
Db      1029 VEITWGINAGKKITNI--QSGDITQNSNDATVGRVVDLKTLESKNSAKTAQNSL 1086

Qy      384 KEFGKLPEDGKDNDFRIKR-YISKYTTINPALTHGVSEYIGSEVGKGLADLVVWNPAPF-- 440
Db      1087 HEFVSVADEQG--N-HFTVSNPY-SSY--DTSKTSVDITPAG--ENG-IITKV--NKGVVR 1135

Qy      441 PCV-----K-----PKIVI---KG-GMVVFSEMGD-----SN--ASVETPQPVYREMGH- 480
Db      1136 VGIDQTKGLTTPKLTGVGNNGKGIVIDSKDGQNTITGLSNTLANV-T-----NDGAGHA 1188

Qy      481 --HGKAKPDT-----SITFVSKVAE---NGVEKGLGLERQVLVFNCRNITKKDP-KF 528
Db      1189 LSQGLAN-DTDTKTRAASIGDVLNAGFNLCNG--EAVDF-----VSTDVTV---DFIDG 1236

Qy      529 NDKTAKITVD--PKT-----FEVFDGKLTCKTSKPTSQVPLAQRYT 566
Db      1237 NATTAKVTVYDTSKTSKVYDVNDNK--TIEVTSKGLGVKTT 1278

RESULT 8
US-10-505-928-98
; Sequence 98, Application US/10505928
```

```
; Publication No. US20060089532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 98
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-98

Query Match      10.5%, Score 315; DB 6; Length 572;
Best Local Similarity 23.4%; Pred. No. 9.6;
Matches 147; Conservative 94; Mismatches 201; Indels 185; Gaps 61;

Qy   55  GQSNSPDNTLD-LVITNAMI--DYTGYYKADIGIKNGKIHGICAGNKQMDGVSPHMV 112
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Qy   56  GKKSIP-HITSRLLIKGRIINDQSLY-ADVYLEGLIKQIGB--N-----LI 50
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy   113 V-GVGTEALAGEGMIIITAGGIDSHTHFLSPQQFTPTALAN---GV-TTFGGGGTGPVD-- 164
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Qy   51  VPG-GVKTIENGWVTPGGIDVNTYLQKPSQGHTA-ADDFQFQTRAAVLGGTTMIIDHV 108
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy   165 ----GNNATTITPKWNHLRMLRAE-----EYSMNVGFGL--KGNSSKKQLVEQVEAG 213
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Qy   109 VPBGSSLLT-SPEKM--H---BAADTKSCDYSLHDVITSWYDGVREELEVLVQ--DKG 160
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy   214 AIGFKL---HED-WGTTPSAIDHCLSVADEVDVOVCIHDTVTNEAGVYDDTLNMNGRAI 269
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Qy   161 VNSFOVMYAKDVYQMSDSOLYAETFLKGLGAVILVHAEB--N-G---D-LIAEQEKRI 212
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy   270 HAYHIE-GAGG--GHS---PD-----VITWAGELN-----ILPSTTPTI---- 303
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Qy   213 ----LENGITGPEGHALSRPEELEAEAVFRATITAGRINCPVIITKVMSKAADIALAR 268
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy   304 ---PYTITNTVAEHLDMLTCHLDKRIRDLOFSQS-RIRPGSIAAEDVLHDMGVIAMT 358
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Qy   269 KGGLPVGEPIASLGTDGT-HYWSKNWAKAAPVTSPPLSDPD-TFPDYLTSL----LA 322
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy   359 SPSQAWGRAGEVIPRWQTADK--NKKBFGKLPE-----D-----GKNDN 398
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Qy   323 CGDLQVTG-SGH-CP--YSTAQAVGKDNTFLIPEGVNGIEERTMVVWDKAVATGKMDEN 378
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy   399 FRIKRVISKYTINPALTHGVSEYIGSVEEGKIADLVWNP-----A-----F 440
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Qy   379 ----QFVAVSTNAAKIFNLPRKRGRIAVGSADADWIWDPDCLKTITPAKSHKSAVEYNIF 434
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy   441 FGVK----PKIVIKGWMVFSEMGDSNAS-----VPT----POPVVYR-----EMFGHHG 482
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Qy   435 EGMECHGSPLVVISQKIVF-EQGINVNNGMGRFIPRKAFPEHLQVRVIRKRVFLQLOG 493
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy   483 --KAKFDTSITFY--SKVAYEN---GVKEKGLGERQVLVPVNCNRNITCKQPFKFNDAKIT 536
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Qy   494 VSRGMYDGPVVEPATPKYATPAPSASKSPS-KHQPPPI---RNLHQSNSFLSG--AQID 547
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy   537 VD---P-KTFEVFV---DGKUCTSKPTS 557
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Qy   548 -DNRPRTGHRIVAPPGR---SNITS 570
Db       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
```

RESULT 9
US-11-301-554-1815
; Sequence 1815, Application US/11301554
; Publication No. US20060098527A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong

```

1  APPLICANT: Watanabe, Yoshihiro
2  APPLICANT: Kalos, Michael D.
3  APPLICANT: Sleath, Paul R.
4  APPLICANT: Johnson, Jeffrey C.
5  APPLICANT: Retter, Marc W.
6  APPLICANT: Durham, Margarita
7  APPLICANT: Carter, Darrick
8  APPLICANT: Fanger, Gary R.
9  APPLICANT: Vedvick, Thomas S.
10 APPLICANT: Bangur, Chaitanya S.
11 APPLICANT: McNabb, Andria
12 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
13 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
14 FILE REFERENCE: 210121.478C21
15 CURRENT APPLICATION NUMBER: US/11/301,554
16 CURRENT FILING DATE: 2005-12-13
17 PRIOR APPLICATION NUMBER: US 10/283,017
18 PRIOR FILING DATE: 2002-10-28
19 PRIOR APPLICATION NUMBER: US 10/113,872
20 PRIOR FILING DATE: 2002-03-28
21 PRIOR APPLICATION NUMBER: US 10/017,754
22 PRIOR FILING DATE: 2001-10-29
23 PRIOR APPLICATION NUMBER: US 09/902,941
24 PRIOR FILING DATE: 2001-07-10
25 PRIOR APPLICATION NUMBER: US 09/849,626
26 PRIOR FILING DATE: 2001-05-03
27 PRIOR APPLICATION NUMBER: US 09/736,457
28 PRIOR FILING DATE: 2000-12-13
29 PRIOR APPLICATION NUMBER: US 09/702,705
30 PRIOR FILING DATE: 2000-10-30
31 PRIOR APPLICATION NUMBER: US 09/677,419
32 PRIOR FILING DATE: 2000-10-06
33 PRIOR APPLICATION NUMBER: US 09/671,325
34 PRIOR FILING DATE: 2000-09-26
35 PRIOR APPLICATION NUMBER: US 09/658,824
36 PRIOR FILING DATE: 2000-09-08
37 Remaining Prior Application data removed - See File Wrapper
38 NUMBER OF SEQ ID NOS: 2157
39 SOFTWARE: FastSeq for Windows Version 4.0
40 SEQ ID NO 1815
41 LENGTH: 572
42 TYPE: PRT
43 ORGANISM: Homo sapiens
44 US-11-301-554-1815

```

Query Match	10.5%;	Score 315;	DB 7;	Length 572;
Best Local Similarity	23.4%;	Pred. No. 9.6;		
Matches	147;	Conservative	94;	Mismatches 201; Indels 185; Gaps 61;
Qy	55	QGSNSPDNTLD-LVITNAMI	-DYTCIYKADIGIKNGKIHGICKAGKMDQGVSPHW	112
Db	5	GKXSIP-HITSRLLIKGRIINDOSLY-ADVULEGLIKQIE-N	-----LI	50
Qy	113	V-GVGTEALAGEGMIITAGIGDSHTHFLSPQQFPTALAN	---GV-TTMFGGGTGPVD-	164
Db	51	VPG-GVKTIENGWVTPGGIDVNTYLYQKSQGWTA-ADOFFQGTAAALVGTTMIIDHV	108	
Qy	165	-----GNATTITPGKNLHRLMLRAAE-----EYSMVNVFLG-	-KGNSSSKQLVEQVEAG	213
Db	109	VPBPGSSLLIT-SPEKW-H---EAAATKSCDYSLHVITDSWYDGVREELEVLYQ--DKG	160	
Qy	214	AIGFKL---HED-WGITPSAIDHCLSLVADEYDVQVCJHTDTVNEAGYVDDTLTNMNGRAI	269	
Db	161	VNSFQVMTAKYQVMSDSQLYEAFTKGLGAVILVHAE--N-G--D-LIAQEKRI	212	
Qy	270	HAYHIE-GAGG--GHS-----PD-----VITMAGELN-----ILPSSSTPTI	303	
Db	213	-----LEWGTGPGSHALSRPEELEAEAVFRAITIAGRINCPVYITKVMKSAAADIILALAR	268	
Qy	304	----PYTI-NTVAEHLDMLTCHHLDKRIREDLOFSQS-RIRPGSIAAEDVLHDMGVAMT	358	
Db	269	KKGFLVFGEP IAAASLGTDGP-HYWSKNWKAARAAAFVTSPLSPDP-TTPDYLTSL-----LA	322	

QY 359 SSDSQMGRAGEVPTWOTADK--NKKEFGKLPB-----D-----GKNDN 398
Db 323 CGDLQVTS--SGH--CP--YSTAQKAVGKDNFTLIPGVNGIERTVTVWDKAVATGKMDN 378
QY 399 FRIKRYISKVTINPALTHGVSEYIGSVBEGKIADLVVWNP-----A-----F 440
Db 379 ----QFVAVTSTNAKFNLFNPRGRTAVGSDADVVDKLTITIAKSHKSAVEYNIP 434
QY 441 FGKVP-----PKIVIKGMVVFSEMGDSNAS-----VPT---PQPVYR-----EMFGHHG 482
Db 435 EGMECHGSPLVVISQKIVF--EDGNINVNKGMGRFIPKAPFEPHLYQKVRKIRNVFGLQG 493
QY 483 --KAKPDTSTIV--SKVAYEN--GVKEKLGLEQVLPVKNCRNITKDKFNDKTAKIT 536
Db 494 VSRGMDPVPVYEPATPKYATPAPSASSPS--KHQPPPI---RNLHOSNFSLSG--AQID 547
QY 537 VD--P-KTFEVFV--DGKLTCKTSKPTS 557
Db 548 -DNNPRGTGHRIVAPPGR---SNITS 570

RESULT 10
US-10-471-571A-3542
; Sequence 3542, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471.571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3542
; LENGTH: 1629
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1629)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-3542

Query Match 10.3%; Score 309; DB 6; Length 1629;
Best Local Similarity 25.6%; Pred. No. 1.7e+02;
Matches 161; Conservative 80; Mismatches 199; Indels 190; Gaps 74;
QY 5 KQEVN--T--YGPT-KGDKV-RLGDT-----DLMAEVE-HDYTTYGEELKFGAGKTIR 51
Db 90 KSEVTSNKDTIEHPFSVKAEDISKEDTPKEADV-AEVQPKSSVTHNAE---TPK-VR 143
QY 52 EGMQSQSPDENTLDLVITNAMIIDYGIYKADIGIKNGKHIGKAGNKMQD---D-G 106
Db 144 ----KARSVDGSEFDITRDSKNVVEST-----PITIQ--GKEHFEG--YGSVDIQKPTDLG 192
QY 107 VSPHVMVGVGTEALAGEGMILTAGID--SHTHPLSPQOFPPTALANGVTTMFGGSGTVPD 164
Db 193 VSEVTRFNVGNES---NGLI---GALQKKNKIDFSKOPNFKVRVANNHQS---NTTG--AD 242
QY 165 GTNATTITPGKWNLRMLRA--ABEYSMMVGLF--KGNSSSKQLVQVEAGAIGFKLHED 222
Db 243 G-----WGF-LPSKGNABEVLITNGILGDKG-----L---VNSG--GPKI--D 277
QY 223 WGTTPSAIDHCLSVADYDVQVCIHTDFTVNBAGYVDDT--LNAMNG--RAIHAYHIEGAGG 280
Db 278 TG-----Y-----IVTSS-----MDKTEKQAGQYRGYGF--VKNDSG 310
QY 281 HSPDVTIMAGELNIPSSTTTPTPTINTVAEHLDMLTCHLDKRIRED--LQF--SQS 336
Db 311 NS-----QMVGE--NI--DKSKTNFLNADNS--TNTSDGKF---H--GQRL--NDVILTYVASTG 358.

QY 337 RIRPOSIAA---EDVLHDMGV-----IAMTSSDSQAMG-RAGEVTPRTWOTADKNKKE 385
Db 359 KMR-ABYAGKTWETSITDLGLSKNQAYNFLTIS--SORWGLNQG--INANGMMRTDLKGE 414
QY 386 FGKLPB-----DGK-DNDNFRIKRYISKVTINPALTHGVSEYIGSVBEGKIADLVVW 436
Db 415 FTFTPEAKTITTELEKKVEEIPFKER---KF--NPDLAG-TEKV--TREGQKGEKTI 466
QY 437 NPAPFVGK-PKIVIKGMVVFSEMGDSNASVPTPOV-----YREMF--GHHGKAKFDT 488
Db 467 TPT---LKNP---LTK--VIISK-GEPKBEI--TKDPINELTEYGPETIAPGH--RDEFPD 514
QY 489 SITFVSK--VAYENGVE-KLG-LER-----QVL---PVQNCRNITKDKFNDKTAKITV 537
Db 515 KLPTGEKEVPGKPIKNPETGTVRPPVDSVTKYQPVKGSIVKEBEIIP--EKERKFN 573
QY 538 D--PKTFEVFVDG---KLCTSKPTSQVPL 561
Db 574 DLAPGTEKTRREGQKGEKTIIT--PTLKNP 602

RESULT 11
US-10-449-902-41462
; Sequence 41462, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-Oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449.902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41462
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-41462

Query Match 10.3%; Score 308; DB 6; Length 1006;
Best Local Similarity 25.5%; Pred. No. 57;
Matches 167; Conservative 89; Mismatches 187; Indels 211; Gaps 77;
QY 10 NTY-GPTKGDKVR--LGDTDLWAE-VEHDYTTYGBELKFGAGK-TIR---E---GMG-Q 56
Db 74 DTFELGAGADTWQGLGD--DAFTDNMIGDYT--GDVIDGLNDYVFPFLSERATGVAF 130
QY 57 SNSPDENTLDLV---ITN--AMIID--YTGIVKADIGIKNGK-----IHGIGKAGNKMQD 105
Db 131 AADPTE-TVTLFGAKLHNEIYQVVDGNTG---AD-NMTGGRWNPFFQNGGADRLDQG 185
QY 106 GYSPHVMVGVGTEAL-AGEGMIITAGGIDSHTHF-----LSPQOFPPTAL--ANGVTM 155
Db 186 G-ADYLSGGAGADVLTGGQGGDYINGGDDKDLFGYGGDNLSSGGNGDDTIEGANGDAL 244
QY 156 FG--G--GTGVPD---GTNATTITPGKWNLRMLRAABEYSM---NVGFL--GKG 198
Db 245 YGEDGNDKLVAGAGN--DYLSSGAGSDTLGG-----AGDDYITGDNKVLGAGSG 294
QY 199 N-----SSSKQL-----VEQVEA-----GAIGFKLH-----EDWGTTPSAIDHCLSV 236
Db 295 NDQILATVNSGKSIDGGSGVDRLELRVYVNGAFTAKAENLVNTLS-D-GTT---IQH--- 346
QY 237 ADEYDVQVCIH-----TDTV-----NEAGY-VDDTLNAMNGR-AIHAYHIEGAG----- 278
Db 347 VEAYDV---HGSNASNQFT--TLGGDDNLGYGGDDTLRAGAGNDSIFA-----GDGRDKL 397

QY 279 -G-CHSPDVTMAGELNLPSSSTPTTPTTINTVAEHLDMLTCHHLDKRIEDLQESQ 335
 Db 398 YGEAGN--DIMPAGQEKDVLGGT-----GIDT-A-YL-----YVD-NYSQSMSPSV 439
 QY 336 SR-----IRPGSIAAB-DVLHDMGVIAMTSSDSOAMGRAGEVIRPW-OTADKNKKEBFGKL 389
 Db 440 SKGKVLNNGTAAATEVFN-----VFTGSGDYL-KAGDAL-QWFFKSGAGNDKLYG--- 490
 QY 390 PEDGKOND--NFRIRK-YISKYINPALTH--GVSEYIGSVBEG--KIADLVVMPAPFG 442
 Db 491 ---KGKGLLDAGIGRDYVNGGAGDDRIDMGDDNTLIG--DGNDV-DITVMEKSSDG 544
 QY 443 VKPKIVIKG--MVVFSE-MGDSNASVTP--QPVYIREMFGHGHKAKFDTISITV-SKV 496
 Db 545 LS-KVSLGAGNDSVEFNAGLGDYIGSLADGSGVDYAEV-----DRSETTKNLSFVLSAN 599
 QY 497 AY-ENGVEKELGLERQVLPVKNCRNI-----TKDKFKENDT-AKITVDPKTFE 543
 Db 600 AHLVNG-----G-AVL--KNFQVHISGGGAD-----DFTGGKLG-D--TFE 636

RESULT 12

US-11-304-590-3
 ; Sequence 3, Application US/11304590
 ; Publication No. US20060159697A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LECLERC, CLAUDE
 ; APPLICANT: EL-AZAMI EL-IDRISSI, MOHAMMED
 ; APPLICANT: LADANT, DANIEL
 ; APPLICANT: BAUCHE, CECILE
 ; APPLICANT: SEBO, PETER
 ; APPLICANT: LOUCKA, JIRINA
 ; APPLICANT: OSICKA, RADIM
 ; TITLE OF INVENTION: MODIFIED BORDETTELLA ADENYLATE CYCLASE COMPRISING OR
 ; TITLE OF INVENTION: LACKING CD18/CD18 INTERACTION DOMAIN AND USES THEREOF
 ; FILE REFERENCE: 2356.0096
 ; CURRENT APPLICATION NUMBER: US/11/304,590
 ; CURRENT FILING DATE: 2005-12-16
 ; PRIOR APPLICATION NUMBER: PCT/EP04/07811
 ; PRIOR FILING DATE: 2004-06-18
 ; PRIOR APPLICATION NUMBER: EP 03291486.3
 ; PRIOR FILING DATE: 2003-06-18
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 3.3
 ; SEQ ID NO 3
 ; LENGTH: 1705
 ; TYPE: PRT
 ; ORGANISM: Bordetella bronchiseptica
 US-11-304-590-3

Query Match 10.3%; Score 308; DB 7; Length 1705;
 Best Local Similarity 24.9%; Pred. No. 2.1e+02;
 Matches 168; Conservative 86; Mismatches 203; Indels 217; Gaps 79;
 QY 3 MKQGEYVNTYPTKDGKVRLLG---DTDLWARVEHDYTYTGBELKFGAGKTIREGQGQSN- 58
 Db 1129 VKNIE--NLHGSSLNDISI-AGDDRDNLWGDGND-TIHGR-----GGDDILRGLGLDTL 1180
 QY 59 -SPDENTL----DLVITNAMIID-YTGIIYKADIGIKNGKIHGKAGNKMDDGVSPHMV 112
 Db 1181 YGEDGNDIFLQDDTETVSD--IDGAGLDTVDY---SAMIHA-GKI-----VAPH-E 1225
 QY 113 VGVGTALAGEGMIITAG--GIDSHTHFLSPQ-QPTALAN---G-VTTMFG-GGTGP 162
 Db 1226 YGFGEIADLSEGVWRKAARGMDYDVSRSVENVIGTSMKDVLLIGDAQNTLMQGGDDT 1285
 QY 163 V---DGNNTATITPGKNLHRLRAAEYSNNVG-----FLKGNSSSKKQLVEQVEAGAI 215
 Db 1286 VRGDDGDD--LLFGDGN--DML-----YG-DAGNDTYLGGIGD-----DTLEGGA- 1326
 QY 216 GFKLHEDW-GTTPSAIDHCLSLVADEYDV-QVCIHTDTV--NEAG-YVDDTLNANGR--- 267

Db 1327 G-----NDWFGQTP-----AREHDVLRGGAGVDTVDYSQAGAHA---GVATGRIGL 1369
 QY 268 AIHAYHIBAGGSHSPDVTMAGELNLPSSSTPTTPTTINTVAEHLDMLTCHHLDKRI 327
 Db 1370 GILA-DL-GAG---RVDKL---GEAG---SSAYDVTSGIENVVGTTEL-----AD-RI 1409
 QY 328 REDLQFSQSRIRPGSIAAEDVL---HDMGVIAMTSSDSOAMGRAGEVIRPW-OTAD-- 380
 Db 1410 TGD---AQNVLIRGAGA-DVLAGEGD-DVLGGGDDQLSGDAGR--DRLYGEAGDWW 1462
 QY 381 --KNKKEFKLPEDGKD-ND--NF-----IKR-YIS--K-Y-----TINPAL 414
 Db 1463 FPDAAANAGNL-LDGDGNDTVDPSGPRGLDAGAKGVFLSLGKGFASLWDEPETS-VL 1520
 QY 415 THGVSEYIGSV-BEGKIAD--LVVMPAPFGVKPKIVIKGGM---VVFSEMGDSNASVPT 468
 Db 1521 RH-IEAVGSRDVLIGDAGANVLN---GLAGNDVLSGAGDDVL---LGDEGSDLLS 1572
 QY 469 POPVYIREMFGHGHKAKFDTISITFVS---KVAYENG---VK-----EKL-----G-- 507
 Db 1573 GD-AGNDLLFGQGD---DTYL-FGAGYGHDTIYESGGGHDITIRINAGADQLMFAROGND 1627
 QY 508 LERQVLPVKNCRNITKCFKENDKTAI-----TVDP-----KTFEV---FVD-GKLC 551
 Db 1628 LEIRILGTDDA--LTVDH-WYRDADHRVEAIHAAQAIDPAGIEKLVEMAQYDPDGA 1684
 QY 552 TSKPTSQVP--LAQ 563
 Db 1685 AAPPAARVPDTLMQ 1698

RESULT 13

US-11-330-403-9773
 ; Sequence 9773, Application US/11330403
 ; Publication No. US20060159563A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark S.
 ; TITLE OF INVENTION: Genes and Uses for Plant Improvement
 ; FILE REFERENCE: 38-21(53629)B
 ; CURRENT APPLICATION NUMBER: US/11/330,403
 ; CURRENT FILING DATE: 2006-01-12
 ; NUMBER OF SEQ ID NOS: 19250
 ; SEQ ID NO 9773
 ; LENGTH: 7362
 ; TYPE: PRT
 ; ORGANISM: Magnetococcus sp. MC-1
 US-11-330-403-9773

Query Match 10.2%; Score 305; DB 7; Length 7362;
 Best Local Similarity 24.9%; Pred. No. 8.6e+03;
 Matches 182; Conservative 93; Mismatches 207; Indels 248; Gaps 82;
 QY 9 VNTYPTKDGKVRLLGDTDL--WAEVEHDY-TTYGEBLK-----FGAKTIRE----- 52
 Db 5010 VNVLGITTSRDVLLGDSGLITW---NSYDITQGAALRSVTSLRDLGGNDTIRTANGDK 5065
 QY 53 ---GMGQ-----SNS-----PDEN-----TLDLVIIT---NAMTIDY 77
 Db 5066 VIVGFGQDSITVQATSNRHHIGDGNVAIAVDTMGITRMETDTVATTGAEDSITIGL 5125
 QY 78 TGIYKADIGIKNGKIHGKAGNKMDDGVSPHMVVGVT--EALAGE-GMI-----IT 128
 Db 5126 DG-DSSDLGI-NVVI-----AGMAD--DTVN---VLGSTTSRDVLLGDNGLITWNSYDIT 5173
 QY 129 AG---GIDSHTHFLSPQF---PT-ALANGVTTMFG---G-----TG-----PVDGTV 167
 Db 5174 QGAALRSVDS-TLYV--ELFGNDTIRTANGDKVIVGGLGQGITVETNSDSKRSQSGDN 5230
 QY 168 AT-TI-TGKNLHRLRAAEYSNN---VGFGLKGNSSSKKQLVEQVEAGAIQFKLHE 221
 Db 5231 VSLTYDTYGBELYLTQLNQKODHNDITITVGI---EDDTST---LGENITG--G--MFD 5281
 QY 222 D---WGTTPS--AI--D--HCLSLVADE---YDVQVCIHTDT-VNEAGYVDDT-LNAMNG 266

Db 5282 DIIKWTSTSEDIFGDNVAILRAQNEASSDQWMLQAKTLVNDKG-GDDTIITSGG 5340
Qy 267 RAI-----HAYHIEGAGGHSPPVITMAGEL-----N-IL-PSSTTPTI--PYTINTV 310
Db 5341 KVLVGGFGADSI SAINGDHL--VIGDNGQDYDKDSENGILRKMQSLDNVIGGEDTI-TL 5397
Qy 311 AEHLDMLTCHHLKRIREDLOFQSRRTPGSI AAEADVLDH-VGIAMTS-SDSQAMGRA 368
Db 5398 REGFVVMGGY-GDEI---TVYTE-----TVAASGVKOLTGIVSVASVSELEAKGRT 5447
Qy 369 GEVTPRTWQTADKNKEP---G-KL-PED-----GKDNDFRIKRVISKYTTIN---P 412
Db 5448 GRYI-----TGDNMEVDLDLHGLYKIOSNDVAVAMGGD-DTISG---NRATTNDLGYO 5498
Qy 413 ALTHGV-SEYIGSVBEGKIA-DLV-----VW-----NPAPF--GVKPKIVIKGMMVFSBMG 460
Db 5499 VLIQMAADTI-TVEDTSTQDVIFGDNVYVERSTSSYFHMINDSLTNQGG-----G 5550
Qy 461 DSNASVPTQPQVYREMPHGHGKAKPDT-----SITFVSKVAYENGVEKGLGLERQVLVK 516
Db 5551 D---TIVTGOGEKW--IVGGMG-A--DTAEVVSINLTDTLFA-GDSVDITFSR-----AK 5597
Qy 517 --NCRNITKKDFE--NDKTKAITVDPKTFEV-FVDGK---L---CTSKPTSQV----- 559
Db 5598 TGNLTSITSTGLYFGGND--ALVVGDG---DVTFTGGKDNDSLIVDSTGATTAEVLAGTA 5652
Qy 560 --PLAQRYTF 567
Db 5653 GDDPNQRFAP 5662

RESULT 14

US-10-471-571A-5160
; Sequence 5160, Application US/10471571A
; Publication No. US20060115490A1

GENERAL INFORMATION:

; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0

; CURRENT APPLICATION NUMBER: US/10/471,571A

; CURRENT FILING DATE: 2003-09-12

; PRIOR APPLICATION NUMBER: GB-0107661.1

; PRIOR FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 5642

; SOFTWARE: SeqWin99, version 1.03

; SEQ ID NO 5160

; LENGTH: 801

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

; FEATURE:

; NAME/KEY: MISC FEATURE

; LOCATION: (1)..(801)

; OTHER INFORMATION: nitrite reductase

US-10-471-571A-5160

Query Match 10.1%; Score 303; DB 6; Length 801;

Best Local Similarity 25.5%; Pred. No. 43;

Matches 171; Conservative 83; Mismatches 172; Indels 244; Gaps 84;

Qy 9 VNTYGPTKGDKVRLGDDTLWA--E-VEHYDTTYG--EBELKGAGK-----TIREGMGQS 57

Db 255 VNR-GIVVNDVMQTSNVAVGECAEHNGKVYGLVAPL-YEQGKVLADHLTKNETNGYK 312

Qy 58 NSPDNTL-----DL-----VITNMI--ID-YGI-----YKADIGIKNGKIHGICA--G 99

Db 313 GSTTFTSLKVGSCDLYSAGQIVENAEIKGIEIFNSVDNNYK-KIFLKGDNV--VGVAVLYG 369

Qy 100 NKDMODGVSP--HMV-VGVGTE--ALAGEGMIIITAGGIDSHHSHFLSPQGFPTA-LANGVT 153

Db 370 --DIDDD--SRFNNMKKGESTEDYTLVS---LDTKGGEAS---LS-----IADWADDET 415

Qy 154 TWFGGGTGPVD-GT--NATTITP-GKNLHRLRAAEYSNMVFLGK-GNSSSK-KQLV 207

Db 416 IC--CGNG-VDKGTIVNA--ITENG-----FTTVEEVTAKT-----KAGNSCKCKPQI 459
Qy 208 EQVEAGAIGFKLHEDW-GTTPSAIDHCLSA-DEYDVQVCIIHTDTTNEAGYVDDTLNAMN 265
Db 460 AQILQHTLG---DDFVAAKPAGICGCTDLTRDQ-----IVTQ-IRAKG-----LK--T 501
Qy 266 GRAI-HAYHIEGAGGHSPPVITMAGELNLPSSSTPTTPTTINTVAEHLDMLTCHHL 324
Db 502 SKEVRHVLNFKNKG-----CP-KCRPAINYILNMVYPH-D-----HED 538
Qy 325 KR-----IREDLQFS---QSR-----IRPGSIAAEDVLHDMGVIAMTSS 360
Db 539 ERESFANERHANLQNDGTFSVIPQMRGCVTDADQLIRLGEVAKK--YH-VPLVKVTG- 594
Qy 361 DSQAMGRAG---EVIPRTWQTAD---KNKKEFKGLPED-----GKDNDF----- 399
Db 595 -SORVGLYGVKKEELPNWE--DLGMRASASAYGKTRSKVSCVGEFCRFGTQVTRLGI 651
Qy 400 RIKRYISKYTTINPALTH-----GVSEYIG--S-VERGKIAD--LV-VWNPAPFGVFKIV 448
Db 652 RLEKTF-EYIDTF--HKFKMGVS--GCPSCVESG-VKDFGIIISVEN-GF-----QIY 697
Qy 449 I--KGMVVFSEMGDSNASVPTQPV-----YREMFHGHGKAKPDTSTITFSKVA 497
Db 698 IGGNGGTEV--EKAEPLTITVETEDEVIKLGCALMQYRET-GIYAE-----TA-PWLRLRG 750
Qy 498 YENGVEKELGL--ERQVLFPVKNCRNITKKDF-KFNDKTKAITVDPKTFEYFVDGKLCISK 554
Db 751 FEN-VKEVL-LDPERQ-----N--EL-----FERIMD--AKKAVEAPWEAI-----TS- 788
Qy 555 PTSQVPLAQR 564
Db 789 -NAQ-----ARR 794

RESULT 15

US-10-505-928-827

; Sequence 827, Application US/10505928

; Publication No. US20060088532A1

GENERAL INFORMATION:

; APPLICANT: Ludwig Institute for Cancer Research et al.

; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES

; FILE REFERENCE: 28967/39178

; CURRENT APPLICATION NUMBER: US/10/505,928

; CURRENT FILING DATE: 2004-08-27

; PRIOR APPLICATION NUMBER: US 60/363,019

; PRIOR FILING DATE: 2002-03-07

; NUMBER OF SEQ ID NOS: 866

; SOFTWARE: PatentIn 3.2

; SEQ ID NO 827

; LENGTH: 4834

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-505-928-827.

Query Match

10.1%; Score 303; DB 6; Length 4834;

Best Local Similarity 22.9%; Pred. No. 3.5e+03;

Matches 157; Conservative 117; Mismatches 212; Indels 201; Gaps 85;

Qy 8 YNVTYGPTKGDKVRLGDDTD-----LWAEVHDY-----TTYGEEELKFG 45

Db 4007 YATGYG--AGGRLIGGTTESVSTPTLLESIQHFVFIKKVAVNSGGKHCALSSSEGEVYSWG 4064

Qy 46 AGKTIREG-MGOSN-SP-DE-NTLDLIVITNMIIDYT--GIYKADI---G-IKN-GKIHG 94

Db 4065 EAE---DGKLGHNRSPCDRPRVIE-SLRGIEVVDVAAGGAHSACTAAGDLTYWGK--G 4118

Qy 95 -IGKAGNKDMQDGVSPHVMVVGTEALAGEGMI-ITAGGIDSHHSHFLSPQGFPTALANGV 152

Db 4119 RYGRLGHSDEQLKPKLV-----EALQHRVVDIACGSDAQTLCLTDDDTVMSWGDGD 4173

Qy 153 TTMFG-GGT-G---P--VD---GTNATTITPG-KWNLHRLRAAEYSNMVNG-F--LGKG 198

Db 4174 YGKLGSGSDCKVPMKIDSITGLGVKVEGSGQFSV-ALTKSGAVYTWGKGDYHRLGHG 4232
Qy 199 NSSKKQLVEQVEAGAIGFKLHEDWGTTTPSAIDHCLSVADYDVOVCJHT--DTVNEAGY 256
Db 4233 SDDHVR--PROVQ-GLQCKKVIAT--ATGSL--HCVCCTE--DGEV--YTWGD--NDEGQ 4280
Qy 257 V-DDTLNAM-NGRAIHAYHIEG-----AGGHSFDPVITMAGELNLIPLSST--TPT-IP 304
Db 4281 LGDGTNAIORPRLVAA--LQCKKNRVACCSAH-----TLWSTS--KPASAGKLPAQVP 4332
Qy 305 YTINTVAE---HL-DMLMTCCHLDKR---I-REDLOFS--OSRIRPGSIAAEDVLHDM 352
Db 4333 MEYNHLQEIPIIALNRNLLHLLHLSLFCPCIPMFDELSLDETGLGP-SVGF-DTLR-- 4388
Qy 353 GVIAWTSDS-----QA-MGR---AGEVIP--RTWQTADKNKKEFGKL--PEDGKNDN 398
Db 4389 GILISQCKEAAFRKVVQATWVRDRQHGFPVELNRI-QV--KRSRSKGLAGP-DG-TKSV 4443
Qy 399 F-RIKRYISKYTNP---ALTHGV--SEYIG-SVEE---G---KIADLV--VWNPAPFGV 443
Db 4444 FGQCAKMSF--GPDLSLLPHRYVKVKFVGESVDDCGGYSSESIAETCBELQN----GL 4497
Qy 444 KPKIVI-KGGWVFESEMGDSNASV---PTPQ-PVYVREMEGHGKAKFDTSTITFVSKVA 497
Db 4498 TPLLIVTPNR---DESG-ANRDCYLLSPAARAPV-HSSMF-----RF---LGVLLGIA 4543
Qy 498 YENGVEKELGL-E---RQV---LPVKNCRNITKKDF-----KFNDKTAKITVDPKTF 542
Db 4544 IRTGSPLSLNLAEFPVKQLAGMSLTADLSEV-DKDFIPGLMYIRDNEATS-----EEF 4596
Qy 543 EVFVDGKCTSKPTS--Q-VPLAQRYT 566
Db 4597 EAM---SLPPTVPSASQDIQLSSKHT 4620

Search completed: August 9, 2006, 21:48:34
Job time : 36 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:38:07 ; Search time 41 Seconds
(without alignment)
1332.954 Million cell updates/sec

Title: US-09-904-994b-3

Perfect score: 2999

Sequence: 1 MKMKQYVNTYGTGDKV.....KLCTSKPTSQVPLAQRYYTF 568

Scoring table:

BLOSUM62
Gapop 4.0 , Gapext 1.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 80.*
2: pir1.*
3: pir2.*
4: pir3.*
5: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2297	76.6	569	1 URKCBP	urease (EC 3.5.1.5)
2	2244	74.8	569	1 S35291	urease (EC 3.5.1.5)
3	2044	68.2	569	2 C36950	urease (EC 3.5.1.5)
4	2027	67.6	571	2 F83681	urease (alpha subu
5	1998	66.6	569	2 D69729	urease (alpha subu
6	1942	64.8	569	2 S75169	urease (EC 3.5.1.5)
7	1916	63.9	568	2 AG2264	urease (EC 3.5.1.5)
8	1900	63.4	569	2 S47104	urease (EC 3.5.1.5)
9	1849	61.7	568	2 C75586	urease, alpha subu
10	1843	61.5	572	2 H64075	urease (EC 3.5.1.5)
11	1838	61.3	566	2 H83037	urease (alpha subu
12	1838	61.3	570	2 S42607	urease (EC 3.5.1.5)
13	1834	61.2	567	1 C36138	urease (EC 3.5.1.5)
14	1826	60.9	567	1 D43719	urease (EC 3.5.1.5)
15	1822	60.8	567	1 S08480	urease (EC 3.5.1.5)
16	1818	60.6	568	2 G85654	probable urease st
17	1818	60.6	568	2 D90794	urease (alpha subu
18	1814	60.5	569	2 A97648	urease (alpha chain
19	1814	60.5	569	2 AG2871	urease (alpha subu
20	1805	60.2	838	2 A96899	probable urease F1
21	1801	60.1	571	2 AF3458	urease (EC 3.5.1.5)
22	1797	59.9	571	2 E90027	urease (alpha subu
23	1795	59.9	572	2 AC0325	urease (EC 3.5.1.5)
24	1792	59.8	571	2 S38485	urease (EC 3.5.1.5)
25	1770	59.0	840	1 URJB	urease (EC 3.5.1.5)
26	1734	57.8	573	2 AI3332	urease (EC 3.5.1.5)
27	1733	57.8	835	2 T37939	urease (EC 3.5.1.5)
28	1721	57.4	598	2 G82890	urease complex com
29	1717	57.3	573	2 S36028	urease (EC 3.5.1.5)

30	1624	54.2	577	2 B70665	probable ureC prot
31	1549	51.7	614	1 S10032	urease (EC 3.5.1.5)
32	1266	42.2	558	2 T29056	urease (EC 3.5.1.5)
33	1222	40.7	569	2 T50711	urease (EC 3.5.1.5)
34	441	14.7	129	2 T08993	urease (EC 3.5.1.5)
35	339	11.3	961	2 AD0548	putative autotransp
36	333	11.1	1078	2 F97907	hyaluronate lyase
37	332	11.1	424	2 H96963	dihydroorotase [im
38	331	11.0	1066	2 B95037	hyaluronidase [imp
39	329	11.0	1983	2 G86643	hypothetical prote
40	328	10.9	579	2 AD1664	adenine deaminases
41	326	10.9	897	2 AD2906	aconitate hydratase
42	326	10.9	897	2 E97681	aconitate hydratase
43	325	10.8	526	2 B75202	d-aminoacylase (as
44	324	10.8	2020	2 C48399	ABC-type transport
45	324	10.8	2198	2 T20371	hypothetical prote

ALIGNMENTS

RESULT 1

URKCBP

urease (EC 3.5.1.5) 62K chain - Helicobacter pylori (strains 26695, J99, and others)
N:Alternate names: urease beta chain; urease chain B; urease large subunit
C:Species: Helicobacter pylori
C:Date: 30-Sep-1991 #sequence revision 02-Dec-1994 #text change 05-Oct-2004
C:Accession: B38537; H64528; A71977; F41834; S07885; S12487; A49215; B61371; A35306; B41:
R:Labigne, A.; Cussac, V.; Courcoux, P.
J. Bacteriol. 173, 1920-1931, 1991

A:Title: Shuttle cloning and nucleotide sequences of Helicobacter pylori genes responsibl
A:Reference number: A38537; MUID:91161505; PMID:2001995
A:Accession: B38537
A:Molecule type: DNA

A:Residues: 1-569 <LAB>

A:Cross-references: UNIPROT:P14917; UNIPARC:UPI000002C20A; GB:M60398; NID:g149007; PIDN:?
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: H64528

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-569 <TOM>

A:Cross-references: UNIPARC:UPI000002C20A; GB:AE000529; GB:AE000511; NID:g2313152; PIDN:?
A:Experimental source: strain 26695
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; J
Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: A71977

A:Molecule type: DNA

A:Residues: 1-569 <ARN>

A:Cross-references: UNIPARC:UPI000002C20A; GB:AE001446; GB:AE001439; NID:g4154573; PIDN:?
A:Experimental source: strain J99
R:Cussac, V.; Ferrero, R.L.; Labigne, A.
J. Bacteriol. 174, 2466-2473, 1992

A:Title: Expression of Helicobacter pylori urease genes in Escherichia coli grown under i
A:Reference number: A41834; MUID:92210488; PMID:1313413

A:Accession: F41834

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 565-569 <CUS>

A:Cross-references: UNIPARC:UPI000002D2AD; GB:M84338
R:Clayton, C.L.; Pallen, M.J.; Kleanthous, H.; Wren, B.W.; Tabaqchali, S.
Nucleic Acids Res. 18, 362, 1990

A:Title: Nucleotide sequence of two genes from Helicobacter pylori encoding for urease st
A:Reference number: S07884; MUID:90221820; PMID:2326167

A:Accession: S07885

A:Molecule type: DNA
A:Residues: 1-9,'A',11-18,'A',20-103,'T',105-180,'F',182-192,'F',194-217,'L',219-272,'Y'
A:Cross-references: UNIPARC:UPI0000172DA6; EMBL:X17079
R:Clayton, C.L.
submitted to the EMBL Data Library, October 1989
A:Reference number: S12487
A:Accession: S12487
A:Molecule type: DNA
A:Residues: 1-9,'A',11-103,'T',105-180,'F',182-192,'F',194-217,'L',219-272,'Y',274-539,'
A:Cross-references: UNIPARC:UPI000002DB18; EMBL:X17079; NID:g43633; PIDN:CAA34933.1; PID
R:Turbett, G.R.; Hoj, P.B.; Horne, R.; Mee, B.J.
Infect. Immun. 60, 5259-5266, 1992
A:Title: Purification and characterization of the urease enzymes of Helicobacter species
A:Reference number: A49215; MUID:93084378; PMID:1452359
A:Accession: A49215
A>Status: preliminary
A:Molecule type: Protein
A:Residues: 1-12 <TUP>
A:Cross-references: UNIPARC:UPI00000B8645; PIDN:AAB24355.1; PID:g260964
A:Experimental source: strain NCTC 11637
R:Evans Jr., D.J.; Evans, D.G.; Kirkpatrick, S.S.; Graham, D.Y.
Microb. Pathog. 10, 15-26, 1991
A:Title: Characterization of the Helicobacter pylori urease and purification of its subu
A:Reference number: A61371; MUID:9312104; PMID:1857197
A:Accession: B61371
A:Molecule type: Protein
A:Residues: 1-21,'C',23-26,'C',28,'I',30 <EVA>
A:Cross-references: UNIPARC:UPI0000172DA7
R:Dunn, B.E.; Campbell, G.P.; Perez-Perez, G.I.; Blaser, M.J.
J. Biol. Chem. 265, 9464-9469, 1990
A:Title: Purification and characterization of urease from Helicobacter pylori.
A:Reference number: A35306; MUID:90264448; PMID:2188975
A:Accession: A35306
A:Molecule type: Protein
A:Residues: 1-20 <DUN>
A:Cross-references: UNIPARC:UPI00000B3DE0
R:Hu, L.T.; Mobley, H.L.T.
Infect. Immun. 58, 992-998, 1990
A:Title: Purification and N-terminal analysis of urease from Helicobacter pylori.
A:Reference number: A41502; MUID:90202165; PMID:2318539
A:Accession: B41502
A:Molecule type: Protein
A:Residues: 1-5,'A',7-15 <HUA>
A:Cross-references: UNIPARC:UPI0000172DA8
C:Genetics:
A:Gene: ureB; HP0072
C:Complex: heterodimer of 26K chain (see PIR:URKCAP) and 62K chain
C:Function:
A:Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and tw
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: heterodimer; hydrolase; metalloprotein; nickel
F:4-552/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted
F:219,248,274/Binding site: nickel 1 (Lys, His, His) #status predicted
F:219/Binding site: carbon dioxide (Lys) (covalent) (by urease activase) #status predict
F:221,322/Active site: His #status predicted
Query Match 76.6%; Score 2297; DB 1; Length 569;
Best Local Similarity 73.4%; Pred. No. 1.9e-46;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
QY 2 KMKQEVNTYPTGDKVRLGDTDLAEVHDYTTTGELKFGAGKTIREGMQSNPSD 61
DB 3 KISRKEYVMTGPTGDKVRLGDTDLAEVHDYTTTGELKFGAGKTIREGMQSNPS 62
QY 62 ENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMQDGVSPHVVGVGTEALA 121
DB 63 KEELDLITNALIVDYTYGIYKADIGIKGKIAGIKGNGKQMDGVKNLSVGPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQOFPFALANGVTTMTFGGGTGPVDGNTATTITPGKNLHRM 181
DB 123 GEGLIIVTAGGIDTHIFHSPOQIPTAFASGVTTMTGGGTGPADGTNATTITPGRNLRKM 182

QY 182 LRAAEYSNNVGFGLKGNSSSKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADEYD 241
DB 183 LRAAEYSNNLGLAKGNASNDASLADQTEAGAIGFKIHEDWGTTPSAINHALVDADKYD 242
QY 242 VQVCIHDTTVNEAGYVDDTLNAMNGRAHYHIEGAGGSHSPDVTIMAGELNILPSTTP 301
DB 243 VQVALHTDTLNEAGCVEDTMAAIGRTMTFTTEGAGGHHAPDIIIKVAGHNLIPASTNP 302
QY 302 TTPYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEVDLHDMGVIAMTSSD 361
DB 303 TTPFTVNTAEAEHMDMLVCHHLDKSKEDVQADSRIRPQTIAAETDLHDMGIFSITSSD 362
QY 362 SQAMGRAGEVIRPTWTQADNKKKFKGKPEDGKDNDFRIKRYISKYTTINPALTHGVSEY 421
DB 363 SQAMGRVGEVIRPTWTQADNKKKFKGKPEDGKDNDFRIKRYISKYTTINPALTHGVSEY 422
QY 422 IGSVEEGKADLAVNPPAFPGVKPIVTKGMVVESEMGSNASVPTPTQVYVREMGHH 481
DB 423 VGSVEVGKADLVLWSPAPFGVKPNMIHKGGFIALSQMGDANASIPTPQVYVREMPFAHH 482
QY 482 GKAKPDTSTITFVSKVAYENGVEKELGLERQVLVKNCRNITKKDKFNDKTKAKITVDPKT 541
DB 483 GKAKYDANITFVSOAYDKGIEKELGLERQVLVKNCRNITKKDMQFNDTTHAIEVNPET 542
QY 542 FEVFDGKLCTSKPTSQVPLAQRYTFF 568
DB 543 YHVFVDGKGVTSKPAKVSQAQLFSIF 569
RESULT 2
S35291
Urease (EC 3.5.1.5) 62K chain - Helicobacter felis
N:Alternate names: urease beta chain
C:Species: Helicobacter felis
C:Date: 31-Dec-1993 #sequence_revision 02-Dec-1994 #text_change 05-Oct-2004
C:Accession: S35291
Mol. Microbiol. 9, 323-333, 1993
R:Ferrero, R.L.; Labigne, A.
A:Title: Cloning, expression and sequencing of Helicobacter felis urease genes.
A:Reference number: S35290; MUID:94018627; PMID:8412683
A:Accession: S35291
A:Molecule type: DNA
A:Residues: 1-569 <PER>
A:Cross-references: UNIPROT:Q08716; UNIPARC:UPI000002F34B; EMBL:X69080; NID:g396160; PIDN
C:Genetics:
A:Gene: ureB
C:Complex: heterodimer
C:Function:
A:Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and tw
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: heterodimer; hydrolase; metalloprotein; nickel
F:1-569/Product: urease 62K chain homology <U62>
F:4-552/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted
F:219,248,274/Binding site: nickel 1 (Lys, His, His) #status predicted
F:219/Binding site: carbon dioxide (Lys) (covalent) (by urease activase) #status predict
F:221,322/Active site: His #status predicted
Query Match 74.8%; Score 2244; DB 1; Length 569;
Best Local Similarity 72.5%; Pred. No. 3.5e-45;
Matches 411; Conservative 71; Mismatches 85; Indels 0; Gaps 0;
QY 2 KMKQEVNTYPTGDKVRLGDTDLAEVHDYTTTGELKFGAGKTIREGMQSNPSD 61
DB 3 KISRKEYVMTGPTGDKVRLGDTDLAEVHDYTTTGELKFGAGKTIREGMQSNPS 62
QY 62 ENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMQDGVSPHVVGVGTEALA 121
DB 63 SYELDLVLTNALIVDYTYGIYKADIGIKGKIAGIKGNGKQMDGVNNLVCVPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQOFPFALANGVTTMTFGGGTGPVDGNTATTITPGKNLHRM 181
DB 123 GEGLIIVTAGGIDTHIFHSPOQIPTAFASGVTTMTGGGTGPADGTNATTITPGRNLRKM 182

Db 123 AEGLVITAGGIDTHIFISPOQIPTAFASGVTTMIGGCTGADGTNATTITPGRANLXSM 182
Qy 182 LRAAEYSMVVGLCKGNSSKKQLVEAGATGFKLHEDWGTTPSAIDHCLSLVADEVD 241
Db 183 LRAAEYAMNGLFLAKGNVSEPSLRDQIEAGATGFKLHEDWGTTPAAIHCLNVADEVD 242
Qy 242 VQVCITHTVNEAGVDDTLNAMGRAIHAHYIEGAGGCHSPDVTIMAGELNIPSSSTTP 301
Db 243 VQVAIHDTLNEAGVDDTLNAMGRAIHAHYIEGAGGCHSPDVTIMAGELNIPSSSTTP 302
Qy 302 TTPYTINTVAEHLDMLTCHHLDRKIREDLQFSOSRIRPGSIAAEDVLHDMGVIAMTSSD 361
Db 303 TTPKTNTAEAEHMDLVMVCHLDRKIREDLQFSOSRIRPGSIAAEDVLHDMGVIAMTSSD 362
Qy 362 SQAMGRAGEVTPRTWQTADKNKKEFGKLPEDGKNDNFRPKRYISKYITINPALTHGVSEY 421
Db 363 SQAMGRAGEVTPRTWQTADKNKKEFGKLPEDGKNDNFRPKRYISKYITINPALTHGVSEY 422
Qy 422 IGSVEEGKIADLVVNNPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFGHH 481
Db 423 VGSVEVGKIADLVVNNPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFGHH 482
Qy 482 GKAKFDTSIITFVSKVAVYGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFGHH 541
Db 483 GKAKFDTSIITFVSKVAVYGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFGHH 542
Qy 542 FEVVDGKLTCTSKPTSQVPLAQRVTFP 568
Db 543 YKVKVKGKVTSKAABELSLAQLNLF 569

RESULT 3
C36950
N;Alternate names: ureC protein
C;Species: Bacillus sp.
C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 05-Oct-2004
R;Maeda, M.; Hidaka, M.; Nakamura, A.; Masaki, H.; Uozumi, T.
J. Bacteriol. 176, 432-442, 1994
A;Title: Cloning, sequencing, and expression of the thermophilic Bacillus sp. strain TB-90
A;Reference number: A36950; MUID:94117379; PMID:8288539
A;Accession: C36950
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-569 <MAE>
A;Cross-references: UNIPROT:Q07397; UNIPARC:UPI0000137D67; GB:D14439; NID:G393296; PIDN:
C;Superfamily: urease, alpha subunit; urease 62K chain homology
F;5-552/Domain: urease 62K chain homology <U62>

Query Match 68.2%; Score 2044; DB 2; Length 569;
Best Local Similarity 66.4%; Pred. No. 2.2e-40;
Matches 378; Conservative 76; Mismatches 108; Indels 7; Gaps 4;

Qy 3 MKKQEVNTYPTKGDVRLDGLDGLAEVEHDYTYGEEELKFGAGKTIREGMQ---SNS 59
Db 5 MSRKQYADMGPTVGDALRLADSLFIEIKDYTYGDEVKFGGKVRDGMGQHPPLATS 64
Qy 60 PDENTLDTVTNAMIIDYTYGKADIGIKNGKIHGIGKAGNKMDDGVSPHVMVGVGTEA 119
Db 65 -DB-CVDLVLNTAIVDVTGYKADIGIKNGKIHGIGKAGNPLMDGV--WVIGAAATEV 120
Qy 120 LAGEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTFGGGTGPVDGTNATTITPGKNLH 179
Db 121 IAAEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTFGGGTGPVDGTNATTITPGKNLH 180
Qy 180 RMLRAAEYSMNVTGFLCKGNSSKKQLVEAGATGFKLHEDWGTTPSAIDHCLSLVADE 239
Db 181 RMLQAAEFINLGLCKGNCSDEAPLKEQIEAGAVGLKHEWGSTAAADITCLKVADR 240
Qy 240 YDQVCHITHTVNEAGVDDTLNAMGRAIHAHYIEGAGGCHSPDVTIMAGELNIPSSST 299

Db 241 YDQVAIHDTLNEGGFVEDTLKAIDRVHTIYHTGAGGGHAPDIIKAAAGFPNIPSSST 300
Qy 300 TTPYTINTVAEHLDMLTCHHLDRKIREDLQFSOSRIRPGSIAAEDVLHDMGVIAMTS 359
Db 301 NTRPTPTINTVAEHLDMLTCHHLDRKIREDLQFSOSRIRPGSIAAEDVLHDMGVIAMTS 360
Qy 360 SDSQAMGRAGEVTPRTWQTADKNKKEFGKLPEDGKNDNFRPKRYISKYITINPALTHGV 419
Db 361 SDSQAMGRAGEVTPRTWQTADKNKKEFGKLPEDGKNDNFRPKRYISKYITINPALTHGV 420
Qy 420 EYIGSVEEGKIADLVVNNPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFG 479
Db 421 DVGVSVEVGKIADLVVNNPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFG 480
Qy 480 HGGKAKFDTSIITFVSKVAVYGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFG 539
Db 481 AKGDGATQTSITFVSKAAVEKGIHQGLKVKVPHGIRKLTCKOLILNDKTPKIDVDP 540
Qy 540 KTFEVPVVGKLTCTSKPTSQVPLAQRVTFP 568
Db 541 QTYEVKVDGQLVTCBPAPVIPAQRVFLP 569

RESULT 4
P83681
urease alpha subunit ureC [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 05-Oct-2004
C;Accession: F83681
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: F83681
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-571 <STO>
A;Cross-references: UNIPARC:UPI00000C3801; GB:AP001507; GB:BA000004; NID:gl0172612; PIDN:
A;Experimental source: strain C-125
C;Genetics:
A;Gene: ureC
C;Superfamily: urease, alpha subunit; urease 62K chain homology

Query Match 67.6%; Score 2027; DB 2; Length 571;
Best Local Similarity 64.4%; Pred. No. 5.8e-40;
Matches 368; Conservative 88; Mismatches 112; Indels 3; Gaps 2;

Qy 1 MKKQEVNTYPTKGDVRLDGLDGLAEVEHDYTYGEEELKFGAGKTIREGMQ---N 58
Db 1 MKLTRAQHSLYGPVGVKVRDLADTLLEIEKDYTYGDEVKFGGKVRDGMGQSAVY 60
Qy 59 SPDENTLDTVTNAMIIDYTYGKADIGIKNGKIHGIGKAGNKMDDGVSPHVMVGVGTE 118
Db 61 TRDEGVLDLIITNATIIDYTYGKADIGIKNGKIHGIGKAGNPDIMDGVESHMIVGASTE 120
Qy 119 ALAGEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTFGGGTGPVDGTNATTITPGKNL 178
Db 121 ALAGEGLVITAGGIDAHIFISPOQIDVAIASSIITMLGGTGPATGATKATCTPGKWN 180
Qy 179 HRMLRAAEYSMNVTGFLCKGNSSKKQLVEAGATGFKLHEDWGTTPSAIDHCLSLVAD 238
Db 181 ERMLEAADAFVNLGFLCKGNASTPAPLREQIEAGAVGLKHEWGTTPPAARTCLSLVAD 240
Qy 239 EYDQVCHITHTVNEAGVDDTLNAMGRAIHAHYIEGAGGCHSPDVTIMAGELNIPSS 298
Db 241 RMDVQVAIHTDTLNEAGFVEDTIIKAGDRVHTIYHTGAGGGHAPDIIKAAAGFPNIPSS 300
Qy 299 TTPYTINTVAEHLDMLTCHHLDRKIREDLQFSOSRIRPGSIAAEDVLHDMGVIAMT 358
Db 301 TNPTPTPTINTVAEHLDMLTCHHLDRKIREDLQFSOSRIRPGSIAAEDVLHDMGVIAMT 360
Qy 359 SSDSQAMGRAGEVTPRTWQTADKNKKEFGKLPED--GKNDNFRPKRYISKYITINPALTHG 417

Db 302 PTRPYTINTLBEHLDMVMCHLHNRNIPEDVAFABESRIRRETIAAEDILHDLGAFSISS 361
Qy 361 DSQAMGRAGEVLPRTWQTADKNKKEFGKLP-EDGKNDNFRKRYISKYTTINPALTGHVS 419
Db 362 DSQAMGRAGEVLPRTWQTAHKQKQVQGLPGETG-NNDNLRAKRYVAKYTTINPAITHGIS 420
Qy 420 EYIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMF 479
Db 421 BEIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMF 480
Qy 480 HHGKAKFDTSTITFVSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKFNDKTAKITVD 539
Db 481 NYGGAIAATSVTFVSKAAKQDIEKGLSKPLVAVKNIQRLTKEDLKNLYLPHIEVD 540
Qy 540 KTFPEVFDGKCTSKPTSOVPLAQRITFF 568
Db 541 EYEVADGELLTCEPASVLPMAQRYFLF 569

RESULT 7
AG2264
urease alpha chain [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 05-Oct-2004
C:Accession: AG2264
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2264
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-568 <UR>
A:Cross-references: UNIPARC:UPI00000CE887; GB:BA000019; PIDN:BA075369.1; PID:gl7132803;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3670
C:Superfamily: urease, alpha subunit; urease 62K chain homology

Query Match 63.9%; Score 1916; DB 2; Length 568;
Best Local Similarity 61.4%; Pred. No. 2.6e-37;
Matches 350; Conservative 95; Mismatches 117; Indels 8; Gaps 4;

Qy 2 KMKKQEVNTYVPTKGDVRLGDTDLWAEVEHDYTYGEEKFKGAGKTIREGMGQS---N 58
Db 4 RMGRQAYETVPTVGDRIADLTFIQVEQDFTTYGDEVKFGGKVRDGMGQSPIAN 63
Qy 59 SPDENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDDQVSPHVVVGTE 118
Db 64 A--DGAVDLVITNALILDWGVKADIGIKDGKIFKIGKAGNPYIQHVD--IIIGPCTE 119
Qy 119 ALAGEGMIITAGGIDSHTHFLSPQOFTALANGVTMTFGGGTGPVDGNTATTTPGKNL 178
Db 120 ALAGEGMILTAGGIDTHIFHCPOQIEVATSGITTMIGGGTGPATCTTCPPGPNM 179
Qy 179 HRMLRAAEYSNMVGLFGKGNSSKKQVQVEAGATGFKLHEDWGTPSAIDHCLSVAD 238
Db 180 YRMLQADAFPMNLGFLGKGNASQOQLGVQIFAGALGLKLHEDWGTPATIDTCLTVAD 239
Qy 239 EYDVQVCIHTDVTNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTIMAGELNLPSS 298
Db 240 EYDVQVAIHTDVTNEAGFVEDTAAFKNRAIHYHIEGAGGHPADIIVKVGQANVLPSS 299
Qy 299 TTPTIPYNTVAEHLDMVMCHLHNRNIPEDVAFABESRIRRETIAAEDILHDLGAFSMI 359
Db 359 SSSQAMGRAGEVLPRTWQTADKNKKEFGKLPEDGKNDNFRKRYISKYTTINPALTGHV 418
Qy 418 AHEVGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMF 478
Db 420 AHEVGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMF 479

Qy 419 SEYIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMF 478
Db 419 AQTVGSVEAGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMF 478
Qy 479 GHGKAKFDTSTITFVSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKFNDKTAKITVD 538
Db 479 GSPAGARNATSLTFVSOAALEREIPQQLGLRKSAAVAVSGTRQLTKQDMKLDALPHIEVD 538
Qy 539 KTFPEVFDGKCTSKPTSOVPLAQRITFF 568
Db 539 SESYEVRADGELLTCEPASVLPMAQRYFLF 568

RESULT 8
S47104
urease (EC 3.5.1.5) 62K chain - Bacillus pasteurii
N:Alternate names: ureC protein
C:Species: Bacillus pasteurii
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S47104
R:Moersdorf, G.; Weimann, P.; Kaltwasser, H.
submitted to the EMBL Data Library, June 1994
A:Description: Nucleotide sequence of three genes on a urease encoding DNA-fragment from
A:Reference number: S47102
A:Accession: S47104
A:Molecule type: DNA
A:Residues: 1-569 <MOE>
A:Cross-references: UNIPROT:P41020; UNIPARC:UPI0000137D66; EMBL:X78411; NID:G498708; PIDN:
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted

Query Match 63.4%; Score 1900; DB 2; Length 569;
Best Local Similarity 59.1%; Pred. No. 6.4e-37;
Matches 337; Conservative 112; Mismatches 118; Indels 3; Gaps 2;

Qy 1 KMKKQEVNTYVPTKGDVRLGDTDLWAEVEHDYTYGEEKFKGAGKTIREGMGQS-- 58
Db 1 MKINRQQAESYVPTVGDRIADLTDL-GEVERDYTLGDEVNFGGKVRREGMGENTY 59
Qy 59 SPDENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDDQVSPHVVVGTE 118
Db 60 TRTENVDLLLTNALILDYTYGIYKADIGVKGQYIVGIGKGNPDINDGVTPNIVGTATE 119
Qy 119 ALAGEGMIITAGGIDSHTHFLSPQOFTALANGVTMTFGGGTGPVDGNTATTTPGKNL 178
Db 120 VIAAEGKIIVTAGGIDTHVHFINPDQVDVALANGITTLFGGCTGPARGSKATTVPGPWNI 179
Qy 179 HRMLRAAEYSNMVGLFGKGNSSKKQVQVEAGATGFKLHEDWGTPSAIDHCLSVAD 238
Db 180 EKMLKSTEGLPINVGILKGHGSSTAPINEQIDAGAAGLKIHEDMGATPASIDRSUTVAD 239
Qy 239 EYDVQVCIHTDVTNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTIMAGELNLPSS 298
Db 240 EADVQVAIHTDVTNEAGFVEDTAAFKNRAIHYHIEGAGGHPADIMAGHPNVLPS 299
Qy 299 TTPTIPYNTVAEHLDMVMCHLHNRNIPEDVAFABESRIRRETIAAEDILHDLGAFSMI 358
Db 300 TNTPRPVTNTIDHLDMLVMCHLHNRNIPEDVAFABESRIRRETIAAEDILHDLGAFSMI 359
Qy 359 SSSQAMGRAGEVLPRTWQTADKNKKEFGKLPEDGKNDNFRKRYISKYTTINPALTGHV 418
Db 360 STDALAMGRAGEVLPRTWQTADKNKKEFGKLPEDGKNDNFRKRYISKYTTINPALTGHV 419
Qy 419 SEYIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMF 478
Db 420 AHEVGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMF 479
Qy 479 GHGKAKFDTSTITFVSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKFNDKTAKITVD 538

H83037
urease alpha subunit PA4868 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 06-Jan-2003
C;Accession: H83037
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brady, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
-; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:2043737; PMID:10984043
A;Accession: H83037
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-566 <STO>
A;Cross-references: UNIPARC:UPI00000C5E4A; GB:AE004900; GB:AE004091; NID:g9951134; PIDN:
A;Experimental source: strain PA01
C;Genetics:
A;Gene: ureC; PA4868
C;Superfamily: urease, alpha subunit; urease 62K chain homology

Query Match 61.3%; Score 1838; DB 2; Length 566;
Best Local Similarity 59.1%; Pred. No. 1.9e-35;
Matches 337; Conservative 88; Mismatches 139; Indels 6; Gaps 3;

QY 1 MKKKQYVNTYGTGKGVRLGDTDLWAEVEHDYTYGELKFGAGKTIREGQGQNSP 60
DB 1 MKISQAYAMFGPTVGDVRLADTLWIEVERDFTVYGEVFGGKVRIDGCMGQSQLG 60

QY 61 DENTLDLVITNMIIDYTYGADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120
DB 61 AAQVDTVTITNALIDHWGIVKADIGLKDGRIAAGKAGNPDMPQGV--NIAICAGTEVI 118

QY 121 AGEEMIITAGGIDSHTHFLSPQPTALANGVTVMFGGTPVDGTNATITPGKWNLHR 180
DB 119 AGEMLITAGGIDTHIFICPQIEEALMSGVTIMGGTGPAAGTATTCSTGPNHWR 178

QY 181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
DB 179 MLQAADAFPMNIGFTGKGNASLPLEEQVLAGAIGLKLHEDWGSTPAADINCLVEAERH 238

QY 241 DVQVCHTDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPVITWAGELNLPSTT 300
DB 239 DIQVAIHDTLNEGSFVETTLGAFKGTIRHTYTEGAGGHPADIIKACGFANVLPSTN 298

QY 301 PTIPYNTINTVAEHLDMLTCHLDRIRDLQFSQRIKRGSIKRIYISKYINPALTHGVSE 360
DB 299 PTRPFTNTIDEHLDMVCHHLDPALAEVAFESIRRETIAEDILHDLGAFMISS 358

QY 361 DSQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRIYISKYINPALTHGVSE 420
DB 359 DSQAMGRVGEVIRPTWQTADKNKQRLDGDGARDNDFRARRVIAKYITINPAITHGISH 418

QY 421 YIGSVBEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMPGH 480
DB 419 EVGSVEAGKADLVLRPAFFGVKPKLSILKGAIAASLMGDINGSIPTPQPVHYRPMFAS 478

QY 481 HGKAKFDTSTITFVKYAYENGVEKGLERQVLVKNCRNITKKDFKNDKTAKITVDPK 540
DB 479 YAGSRATSLTTFVSOAFAAGVPPQGLRKAIGVVGSGRGVQKTDLHNGLYLPTIEYDAQ 538

QY 541 TPEVFVDGKL--CTSKPTSOVPLAQRVTF 568
DB 539 NYQVRADGQLLWC--EPADVLPMAQRVFLP 566

RESULT 12
S42607
urease (EC 3.5.1.5) - Rhizobium meliloti
C;Species: Rhizobium meliloti
C;Date: 25-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C;Accession: S42607
R;Miksch, G.; Arnold, W.; Lentzsch, P.; Priefer, U.B.; Puehler, A.

Mol. Gen. Genet. 242, 539-550, 1994
A;Title: A 4.6 kb DNA region of Rhizobium meliloti involved in determining urease and hyc
ing frames.
A;Reference number: S42601; MUID:94166766; PMID:8121412
A;Accession: S42607
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-570 <MIK>
A;Cross-references: UNIPROT:P42885; UNIPARC:UPI000016FFCB; GB:S69145; NID:g545795; PIDN:
C;Genetics:
A;Gene: ureC
C;Superfamily: urease, alpha subunit; urease 62K chain homology
C;Keywords: hydrolase; metalloprotein; nickel
F;5-53/Domain: urease 62K chain homology <U62>
F;136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted

Query Match 61.3%; Score 1838; DB 2; Length 570;
Best Local Similarity 59.8%; Pred. No. 2e-35;
Matches 340; Conservative 95; Mismatches 130; Indels 4; Gaps 3;

QY 2 MKKKQYVNTYGTGKGVRLGDTDLWAEVEHDYTYGELKFGAGKTIREGQGQNSPD 61
DB 4 RMSRAAYANMFGPTVGDVRLADTLFIEVKDFTTHGEVFGGKVRIDGCMGQSQVTR 63

QY 62 E-NTLDLVITNMIIDYTYGADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120
DB 64 EGGAVDTVTITNALIDHWGIVKADIGLKDGRIAAGKAGNPDMPQGV--IIIVGPGTEVI 121

QY 121 AGEEMIITAGGIDSHTHFLSPQPTALANGVTVMFGGTPVDGTNATITPGKWNLHR 180
DB 122 AGEKIVTAGMDSHIFICPQIEEALMSGVTIMGGTGPAHGTATTCFPGPHIAR 181

QY 181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
DB 182 MIRAADA FPMNLA PAGKGNASLPALVEMVLGATSLKLHEDWGTTPAALIDCCLSLVADEY 241

QY 241 DVQVCHTDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPVITWAGELNLPSTT 300
DB 242 DVQVMIHTDLNESGFVEDTIAAIGKRTIHAHYTEGAGGHPADIIIRICQPNVIPSSTN 301

QY 301 PTIPYNTINTVAEHLDMLTCHLDRIRDLQFSQRIKRGSIKRIYISKYINPALTHGVSE 360
DB 302 PTRPVTNTLAEHLDMVCHHLSPITPIDIAFASIRIRKETIAEDILHDLGAFSLISS 361

QY 361 DSQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRIYISKYINPALTHGVSE 420
DB 362 DSQAMGRVGEVIRPTWQTADKNKQRLKKEBTGNDNDFRVKRYIAKYITINPAIAGLSH 421

QY 421 YIGSVBEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMPGH 480
DB 422 EIGSLVSGKADLVLRNPAFFGVKPKDMVLLGGTIAAAPMGDPNASIPTPQPVHYRPMFGA 481

QY 481 HGKAKFDTSTITFVKYAYENGVEKGLERQVLVKNCR-NITKKDFKNDKTAKITVDP 539
DB 482 YGRSRTNSSLTTFVSPASLDAGLAGRLGAKELAVQNTRGGIGKASMIHNSLTPHIEVDP 541

QY 540 KTFEYFVDGKLCTSKPTSOVPLAQRVTF 568
DB 542 ETYEVRADGELLTCEPATVLPMAQRVFLP 570

RESULT 13
C36138
urease (EC 3.5.1.5) 62K chain [validated] - Klebsiella pneumoniae
N;Alternate names: urea amidohydrolase; urease alpha chain; urease chain C
C;Species: Klebsiella pneumoniae
C;Date: 30-Nov-1990 #sequence_revision 02-Dec-1994 #text_change 05-Oct-2004
C;Accession: C36138
R;Mulrooney, S. B.; Hausinger, R. P.
J. Bacteriol. 172, 5837-5843, 1990
A;Title: Sequence of the Klebsiella aerogenes urease genes and evidence for accessory pr
A;Reference number: A36138; MUID:91008957; PMID:2211515
A;Accession: C36138

Db	539	QTYIVKADGVPLVC--EPATELPMQARYFLP 567	Db	539	QTYIVKADGVPLVC--EPATELPMQARYFLP 567
RESULT 15					
S08480					
urease (EC 3.5.1.5) 62K chain - Proteus vulgaris					
N:Alternate names: urease alpha chain					
C:Species: Proteus vulgaris					
C>Date: 29-Jan-1993 #sequence_revision 02-Dec-1994 #text_change 05-Oct-2004					
C:Accession: S08480					
R:Moersdorf, G.; Kaltwasser, H.					
FEMS Microbiol. Lett. 66, 67-74, 1991					
A:Title: Cloning of the genes encoding urease from Proteus vulgaris and sequencing of th					
F:134,136,217,360/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted					
F:217,246,272/Binding site: nickel 1 (Lys, His) #status predicted					
F:217/Binding site: carbon dioxide (Lys) (covalent) (by urease activase) #status predict					
F:219,320/Active site: His #status predicted					
Query Match 60.8%; Score 1822; DB 1; Length 567;					
Best Local Similarity 58.8%; Pred. No. 4.8e-35;					
Matches 336; Conservative 95; Mismatches 133; Indels 7; Gaps 4;					
Qy	1	MK-MKKQEVNTYPTKGDVRLGDTDLWAEVHDYTYGEELEKFGACKTIREGMGQNS 59			
Db	1	MKTISQAYADMFGPTTGDRLADTELFLIEQDFTTYGEEVKGKVRDGMGQSQV 60			
Qy	60	PDENTLDVITNAMIIDYTVIGYKADIGIKNGKIGIKGAGKNDQDGVSPHMVVGVTGEA 119			
Db	61	VSACVDVLTITNAILIHDHWGIVKADIGIKGRTIGIKAGNPDPQPNVD--IVIGPGTEV 118			
Qy	120	LAGEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDTNATTITPKKMWLH 179			
Db	119	VAGEGKIITAGGVDTIHIFICPQAAEEGLISGVITFIGGTGPVAGTNATTVPGIWNMH 178			
Qy	180	RMLRAAEYSNNVGFLEKGNSSSKQLVQVEAGAIGPKLHEDWGTTPSAIDHCLSLVADE 239			
Db	179	RMLEAVDELPIVGLFGKVCVQPEAIREQIEAGAIGLKIHEWDGATPMAIHNCINVADE 238			
Qy	240	YDVQVCIHTDTVNEAGYVDDTLNANNGRAIHAYHIEGAGGCHSPDVITMAGELNLPST 299			
Db	239	MDVQVAIHSDTLNEGGFYETVKAIGRVHVFHTEGAGGGHAPDVTKSVGEPNLPAST 298			
Qy	300	TPPTPYTINTVAEHLDMLTCKHLDKRIREDLPFSQSRIRPGSIAAEDVLHDMGVIAMTS 359			
Db	299	NPTMPYTINTVDEHLMVCHLDPSIPEDVFAESRIRRETIAAEDILHDMGAI SVMS 358			
Qy	360	SDSQAMGRAGEVI PRTWQTADKNKFKGLPEDGKNDNFRIKRYISKYITINPALTHGVS 419			
Db	359	SDSQAMGRVGEVWRTWQCAHKMKLQKSLAGDTAENDNNRIKRYIAKYITINPALAHGIA 418			
Qy	420	EYIGSVBEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMFG 479			
Db	419	HEVGSIEKGLADIVLMDPAFFGFKPALIMKGMVAVAPMGDINAAIPTQPVHYRPMYA 478			
Qy	480	HGKAKFDTSTITFVSKVAYENGVEKGLERQVLPVKNCRNITKKDKFNDKTAKITVDP 539			
Db	479	CLGRAKYQTSIMFMSKAGIDNGVPEKGLQLGIRVSGCRKVTKASIMHNSYVPHIELEP 538			
Qy	540	KTFEVFDG--KLCTSKPTSQVPLAORYTFF 568			

Search completed: August 9, 2006, 21:43:46
Job time : 43 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:34:37 ; Search time 305 Seconds

(without alignments)
1722.652 Million cell updates/sec

Title: US-09-904-994B-3

Perfect score: 2999

Sequence: 1 MKMKQEVNTYGTGDKV.....KLCTSKPTSQVPLAQRYTFF 568

Scoring table: BLOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2326	77.6	569	2	Q8GH97_HELBI
2	2304	76.8	568	1	URE2_HELHE
3	2301	76.7	569	2	Q7X3W5_HELPC
4	2297	76.6	569	1	URE2_HELPC
5	2297	76.6	569	1	URE2_HELPC
6	2297	76.6	569	2	Q9S0Q5_HELPC
7	2293	76.5	568	2	Q38Q64_HELPC
8	2292	76.4	567	2	Q64EY3_HELPC
9	2291	76.4	569	2	Q9AFB1_HELPC
10	2288	76.3	569	2	Q84F75_HELPC
11	2251	75.1	559	2	Q8RNU6_HELPC
12	2244	74.8	569	1	URE2_HELFE
13	2109	70.3	569	2	Q93FU4_HELHP
14	2044	68.2	569	1	URE1_BACSB
15	2027	67.6	571	2	Q9KG59_BACHD
16	1999	66.7	569	2	Q5KYM1_GEOKA
17	1998	66.6	569	1	URE1_BACSU
18	1986	66.2	572	2	Q4CHE3_CLOTF
19	1971	65.7	569	2	Q52305_synchococ
20	1942	64.8	569	1	URE1_SYNY3
21	1942	64.8	570	2	Q2ZGT0_CALSA
22	1932	64.4	568	2	Q3M712_ANAVT
23	1932	64.4	570	2	Q733J6_BACCL
24	1930	64.4	572	2	Q8DMV6_SYNEL
25	1929	64.3	568	2	Q3FBA4_9BURK
26	1926	64.2	568	2	Q454M1_9BURK
27	1926	64.2	568	2	Q4LNH4_9BURK
28	1924	64.2	565	2	Q5FB23_CAMELA
29	1921	64.1	566	2	Q48DE6_PSEI4
30	1921	64.1	566	2	Q87VP0_PSESM
31	1920	64.0	566	2	Q4ZN06_PSEU2

32	1920	64.0	568	2	Q39TW3_BURS3
33	1916	63.9	568	2	Q8YQZ0_ANASP
34	1914	63.8	569	2	Q46IY3_PROMT
35	1910	63.7	568	2	Q4BIH6_BURVI
36	1905	63.5	566	2	Q3KIT2_PSEPF
37	1905	63.5	568	2	Q34XU6_9GAMM
38	1905	63.5	568	2	Q63RL3_BURPS
39	1904	63.5	572	2	Q2JQ88_9CVAN
40	1902	63.4	566	2	Q4KJ10_PSEFS
41	1902	63.4	568	2	Q6ZHS0_BURMA
42	1901	63.4	564	2	Q3JFJ6_BURPI
43	1900	63.4	566	2	Q4IXD2_AZQVI
44	1900	63.4	569	1	URE1_BACPA
45	1891	63.1	570	2	Q3WT52_9RHIZ

ALIGNMENTS

RESULT 1
Q8GH97_HELBI PRELIMINARY; PRT; 569 AA.

AC Q8GH97;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.

DT 01-MAR-2003, sequence version 1.

DT 07-FEB-2006, entry version 14.

DE Urease B.

OS Helicobacter bizzozeronii.

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

OC Helicobacteraceae; Helicobacter.

OX NCBI_TaxID=56877;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22540228; PubMed=12652903; DOI=10.1080/1042517021000039230;

RA Zhu J., Teng C.H., Chang C.F., Chang C.D., Simpson K.W., Wei C.,

RA McDonough P., McDonough S., Chang Y.F.

RT "Cloning and characterization of a Helicobacter bizzozeronii urease

gene cluster.";

RL DNA Seq. 13:321-331(2002).

CC Copyright by the UniProt Consortium, see http://www.uniprot.org/terms

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL; AF330621; AA015374.1; -; Genomic_DNA.

DR HSSP; P14917; IE9Y.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0016151; F:nickel ion binding; IEA.

DR GO; GO:0009039; F:urease activity; IEA.

DR GO; GO:0019627; F:urea metabolism; IEA.

DR InterPro; IPR006680; Amidohydro 1.

DR InterPro; IPR011612; Urease_alpha N.

DR InterPro; IPR008295; Urease_alphaIone.

DR Pfam; PF01979; Amidohydro 1.

DR Pfam; PF00449; Urease_alpha; 1.

DR PIRSF; PIRSF001226; Urease_alpha; 1.

DR PRINTS; PR01752; UREASE.

DR TIGRFAMs; TIGR01792; urease_alph; 1.

DR PROSITE; PS01120; UREASE_1; 1.

DR PROSITE; PS00145; UREASE_2; 1.

SQ SEQUENCE 569 AA; 61922 MW; 3938FFC83EC39E26 CRC64;

Query Match 77.6%; Score 2326; DB 2; Length 569;

Best Local Similarity 74.1%; Pred No. 3 6e-43;

Matches 420; Conservative 75; Mismatches 72; Indels 0; Gaps 0;

Qy 2 KMKQEVNTYGTGDKVRLGDTDLMAVEHDYTYTGELKFGAGKTIREGQNSPD 61

3 KISKEYVSMYGTGDKVRLGDTDLILEVHDCTTYGEEKFGGKTIRDGMAQTNPS 62

62 ENTLDLVTNMIIDYGIYKADIGIKNGKIHGICAGKMDQGVSPHMVVGVTGTEALA 121

63 SHELDVLVTNALIVDTYGIYKADIGIKNGKIHGICAGKMDQGVCPNLCVGPATEALA 122

```

Qy 122 GEGMIITAGGIDSHFLSPQOFTALANGVTTMFGGTGPVDTGNATTITPGKWNLRHM 181
Db 123 AEGLVITAGGIDTHIFISPOQITAFASGITIMGGTGPADGTNATTITPGRWNLRKM 182
Qy 182 LRAAEYSMNVLGKGNSSKKOLVQVEAGAGFKLHEDWGTTPSAIDHCLSVADYD 241
Db 183 LRAEEYAMNLGYLGKGNVSEPSLVQLEAGAGFKLHEDWGTTPAAYHCLNVADYD 242
Qy 242 VQVCITDTNEAGYVDDTLNMMNGRAIHYHIEGAGGGHSPDVTWAGELNLPSTTP 301
Db 243 VQVAITDTNEAGCVEDTLQAIAGRTIHTHTTEGAGGGHAPDVIKMSGEFNILPASTNP 302
Qy 302 TIPTVINTVAEHLDMLTCHLDKRIREDLPQSQRIRPGSIAEDVLHDMGVAMTSSD 361
Db 303 TIPTVTNTEAHEMDLVMCHLDKRIREDLPQADSRIRPQTIAAEDKLHDMGIFSTSSD 362
Qy 362 SQAMGRAGEVITPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTINPALTHGVSEY 421
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGRLPEEKGDNDNFRIKRYISKYTINPAIHGISEY 422
Qy 422 IGSVEEKGIALDLVWNPAPFAGKPKIVIKGGMVFSEMGDSNASVPTPQPVYVREMFGHH 481
Db 423 VGSVEVGKPADLVMSAPFAGIKENMIKGGFIALSQMGDANASIPTPQPVYVREMFGHH 482
Qy 482 GKAKFDTSTITFVSKVAYENGVEKLGLEQVLVPVKNCRNITKKDFKNDKTAKITVDPKT 541
Db 483 GKAKFDTNITFVSVQVAYDNGIKKEBLGRLQVLPVKNCRNITKKDLKFNVDVTAHIEVNPET 542
Qy 542 FEVFDGKCLTSKPTSQVLAQRTFF 568
Db 543 YKVKVDGKEVTSKAADKISLAQLNLF 569

RESULT 2
URE2_HELHE
ID_URE2_HELHE STANDARD; PRT; 568 AA.
AC P42823;
DT 01-NOV-1995, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1995, sequence version 1.
DT 07-FEB-2006, entry version 44.
DE Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN Name=ureb;
OS Helicobacter heilmannii.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=35817;
[1]
NP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=2;
RX MEDLINE=94222523; PubMed=8168924;
RA Solnick J.V., O'Rourke J., Lee A., Tompkins L.S.;
RT "Molecular analysis of urease genes from a newly identified uncultured
species of Helicobacter.";
RL Infect. Immun. 62:1631-1638 (1994).
CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -!- COPACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -!- SUBUNIT: Consists of two subunits (alpha and beta).
CC -!- FM: lys-219 is carbamylated. The carbamoyl group provides the
ligands for the two nickel ions (By similarity).
CC -!- SIMILARITY: Belongs to the urease family.
CC -!- CAUTION: In Helicobacter the beta subunit is what is known in
other bacteria as the alpha subunit.
-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL: L25079; AAA65723.1; -; Genomic_DNA.
CC HSSP: P14917; 1E9Z
CC InterPro: IPR006680; Amidohydro 1.
CC InterPro: IPR005848; Urease_alpha.
CC InterPro: IPR011612; Urease_alpha_N.
CC InterPro: IPR008295; Urease_alphalone.

```

```

DR Pfam: PF01979; Amidohydro 1; 1.
DR Pfam: PF00449; Urease_alpha; 1.
DR PIRSF: PIRSF001226; Urease_alpha; 1.
DR PRINTS: PRO1752; UREASE.
DR TIGRFAMs: TIGR01792; urease_alph; 1.
DR PROSITE: PS01120; UREASE_1; 1.
DR PROSITE: PS00145; UREASE_2; 1.
KW Hydrolase; Metal-binding; Nickel.
FT CHAIN 1 568 Urease beta subunit.
FT ACT_SITE 321 321 Proton donor (By similarity).
FT METAL 136 136 Nickel 2 (By similarity).
FT METAL 138 138 Nickel 2 (By similarity).
FT METAL 219 219 Nickel 1 (By similarity).
FT METAL 219 219 Nickel 2 (By similarity).
FT METAL 248 248 Nickel 1 (By similarity).
FT METAL 274 274 Nickel 1 (By similarity).
FT METAL 361 361 Nickel 2 (By similarity).
FT BINDING 221 221 Substrate (By similarity).
FT SEQUENCE 568 AA; 61870 MW; E2C036DF30377917 CRC64;

Query Match 76.8%; Score 2304; DB 1; Length 568;
Best Local Similarity 73.7%; Pred. No. 1.1e-42;
Matches 418; Conservative 74; Mismatches 74; Indels 1; Gaps 1;

Qy 2 KMKQDEYVNTYGTGDKVRLGDTDLMAVEHHDYTYTGSELKPGACKTIREGQNSPD 61
Db 3 KISKEYSVMTGPTGDKVRLGDTDLILEVEHDCITYGSEIKPGGKTIRDCMGQTNPS 62
Qy 62 ENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDQGVSPHVMVGVGTEALA 121
Db 63 SHELDLVITNALIVDYTYGIKADIGIKNGKIHGKAGNKLDQDGVNCLRCVGPTEALA 122
Qy 122 GEGMIITAGGIDSHFLSPQOFTALANGVTTMFGGTGPVDTGNATTITPGKWNLRHM 181
Db 123 AEGLVITAGGIDTHIFISPOQITAFASGITIMGGTGPADGTNATTITPGRWNLRKM 182
Qy 182 LRAAEYSMNVLGKGNSSKKQLVQVEAGAGFKLHEDWGTTPSAIDHCLSVADYD 241
Db 183 LRAEEYAMNLGYLGKGNVSEPPALIDQLEAGAGFKIHEHMGSTPSAINHALNIADKYD 242
Qy 242 VQVCITDTNEAGYVDDTLNMMNGRAIHYHIEGAGGGHSPDVTWAGELNLPSTTP 301
Db 243 VQVAITDTNEAGCVEDTLQAIAGRTIHTHTTEGAGGGHAPDVIKMSGEFNILPASTNP 302
Qy 302 TIPTVINTVAEHLDMLTCHLDKRIREDLPQSQRIRPGSIAEDVLHDMGVAMTSSD 361
Db 303 TIPTVTNTEAHEMDLVMCHLDKRIREDLPQADSRIRPQTIAAEDKLHDMGIFSTSSD 361
Qy 362 SQAMGRAGEVITPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTINPALTHGVSEY 421
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGRLPEEKGDNDNFRIKRYISKYTINPAIHGISEY 421
Qy 422 IGSVEEKGIALDLVWNPAPFAGKPKIVIKGGMVFSEMGDSNASVPTPQPVYVREMFGHH 481
Db 423 VGSVEVGKPADLVMSAPFAGIKENMIKGGFIALSQMGDANASIPTPQPVYVREMFGHH 481
Qy 482 GKAKFDTSTITFVSKVAYENGVEKLGLEQVLVPVKNCRNITKKDFKNDKTAKITVDPKT 541
Db 483 GKAKFDTNITFVSVQVAYENGVEKLGLEQVLVPVKNCRNITKKDLKFNVDVTAHIEVNPET 541
Qy 542 FEVFDGKCLTSKPTSQVLAQRTFF 568
Db 543 YKVKVDGKEVTSKAADKISLAQLNLF 568

RESULT 3
Q7X3W5_HELPY
ID Q7X3W5_HELPY PRELIMINARY; PRT; 569 AA.
AC Q7X3W5;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 12.

```

```
DE Urease B.
GN Name=ureB;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MEL-HP27;
RA Duan G.C., Dai L.P.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR HSSP; AY295085; AAF51176.1; -; Genomic_DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0019627; F:urea metabolism; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR011612; Urease_alpha_N.
DR InterPro; IPR008295; Urease_alpha1one.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
DR SEQUENCE 569 AA; 61671 MW; 4C96852826C7A784 CRC64;
Query Match 76.7%; Score 2301; DB 2; Length 569;
Best Local Similarity 73.5%; Pred No. 1.3e-42;
Matches 417; Conservative 75; Mismatches 75; Indels 0; Gaps 0;
QY 2 KMKKQEVNTYGTGDKVRLGDTDLMAEVEHDYTTTGEELKFGAGKTIREGMQSNPSD 61
DB 3 KISKEVSVMYGPTGDKVRLGDTDLMAEVEHDYTTTGEELKFGAGKTIREGMQSNPS 62
QY 62 ENTLDLVITNAMIIDYTGIVKADIGIKNGKIHGKAGNKMQDGVSPHMVVGVTAL 121
DB 63 KEELDIIITNALIVDYTGIVKADIGIKNGKIHGKAGNKMQDGVSPHMVVGVTAL 122
QY 122 GEGMILITAGGIDSHTHFLSPQEPPTALANGVTYTMFGGGTGPVDCGTNATTITPGKWLHRM 181
DB 123 GEGMILITAGGIDSHTHFLSPQEPPTALANGVTYTMFGGGTGPVDCGTNATTITPGKWLHRM 182
QY 182 LRAAEYSNMVGLFGKGNSSKKQLVQVEAGAIQFKLHEDWGTTPSAIDHCLSVADBYD 241
DB 183 LRAAEYSNMVGLFGKGNSSKKQLVQVEAGAIQFKLHEDWGTTPSAIDHCLSVADBYD 242
QY 242 VQVCIHTDTVNEAGYVDDTLNANMGRAIHAYHIEGAGGSHSPDVTWAGELNLPSSSTP 301
DB 243 VQVCIHTDTVNEAGYVDDTLNANMGRAIHAYHIEGAGGSHSPDVTWAGELNLPSSSTP 302
QY 302 TIPTTINTVAEHLDMLTCHLDRREDLQFSQSRIRPGSIRPAAEDVLDHMGVAMTSSD 361
DB 303 TIPTTINTVAEHLDMLTCHLDRREDLQFSQSRIRPAAEDVLDHMGVAMTSSD 362
QY 362 SQAMGRAGEVIRPTWOTADKNKKEFGKLPEDGKNDNFRIKRYISKYITNPALTHGVSEY 421
DB 363 SQAMGRAGEVIRPTWOTADKNKKEFGKLPEDGKNDNFRIKRYISKYITNPALTHGVSEY 422
QY 422 IGSVEEGKIADLVNMPAFPGVFKPKVITKGMVVFSEMGDSNASVPTPQPVYREMGFH 481
DB 423 VGSVEGKIVADLVNMPAFPGVFKPKVITKGMVVFSEMGDSNASVPTPQPVYREMGFH 482
QY 482 GKAKFTSTITFVSKVAYENGVKELGLERQVLVPGKNCRNITKKDFKNDKTAKITVDPKT 541
DB 483 GKAKYDANITFVSKVAYENGVKELGLERQVLVPGKNCRNITKKDFKNDKTAKITVDPKT 542
```

```
QY 542 FEVFDVGGKLGKTSKPTSQVPLAQRYTFF 568
DB 543 YHFVDGKEVTSKPTKVKSLAQLFISIF 569
RESULT 4
URE2_HELPJ STANDARD; PRT; 569 AA.
ID URE2_HELPJ
AC P69997; P14917;
DT 04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT 04-JAN-2005, sequence version 1.
DT 07-MAR-2006, entry version 7.
DE Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN Name=ureB; Synonyms=hpb; OrderedLocNames=JHP0067;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -!- COFACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -!- SUBUNIT: Heterohexamers of 3 alpha and 3 beta subunits. Four
heterohexamers assemble to form a 16 nm spherical complex (By
similarity).
CC -!- PTM: Lys-219 is carbamylated. The carbamoyl group provides the
ligands for the two nickel ions (By similarity).
CC -!- MISCELLANEOUS: Oligomerization may protect the enzyme against
denaturation in an acidic environment (By similarity).
CC -!- SIMILARITY: Belongs to the urease family.
CC -!- CAUTION: In Helicobacter the beta subunit is what is known in
other bacteria as the alpha subunit.
Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
Distributed under the Creative Commons Attribution-NoDerivs License
EMBL; AE001446; AAD05651.1; -; Genomic_DNA.
PIR; B38537; URKCBP.
GenomeReviews; AE001439; GR; JHP0067.
InterPro; IPR006680; Amidohydro_1.
InterPro; IPR005848; Urease_alpha.
InterPro; IPR011612; Urease_alpha_N.
InterPro; IPR008295; Urease_alpha1one.
Pfam; PF01979; Amidohydro_1; 1.
Pfam; PF00449; Urease_alpha; 1.
PIRSF; PIRSF001226; Urease_alpha; 1.
PRINTS; PR01752; UREASE.
TIGRFAMs; TIGR01792; urease_alph; 1.
PROSITE; PS01120; UREASE_1; 1.
PROSITE; PS00145; UREASE_2; 1.
Complete proteome; Hydrolase; Metal-binding; Nickel.
CHAIN 1 569
Urease beta subunit.
ACT_SITE 322 322 Proton donor (By similarity).
METAL 136 136 Nickel 2 (By similarity).
METAL 138 138 Nickel 2 (By similarity).
FT METAL 219 219 Nickel 1 (By similarity).
FT METAL 219 219 Nickel 2 (By similarity).
FT METAL 248 248 Nickel 1 (By similarity).
FT METAL 274 274 Nickel 1 (By similarity).
FT METAL 362 362 Nickel 2 (By similarity).
FT BINDING 221 221 Substrate (By similarity).
SQ SEQUENCE 569 AA; 61684 MW; 4C8A6BC6C8295584 CRC64;
```

```
Query Match      76.6%; Score 2297; DB 1; Length 569;
Best Local Similarity 73.4%; Pred. No. 1.6e-42;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

QY 2 KMKKEQVNTYPTKDKVRLGDDTLAEVEHDYTYGEEELKFGAGKTIRSGMGQNSPD 61
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 3 KISRKEVSMYPTKDKVRLGDDTLAEVEHDYTYGEEELKFGAGKTIRSGMGQNSPD 62
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 62 ENTLDLVTNAMIIDYTIYKADIGIKNGKIHGIGKAGNKDQDQGVSPHVVGVGTEALA 121
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 63 KEELDILITNALIVDYTIYKADIGIKNGKIHGIGKAGNKDQDQGVKXNLSVGPATEALA 122
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 122 GEGMIITAGGIDSHFLSPQFPPTALANGVYTWFGGCTGPGVDGTNATITPGKWNLRHM 181
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 123 GEGLIVTAGGIDTHIFPISQPIPTAFASGVTTMIGGCTGPGADGTNATITPGRRNLKWM 182
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 182 LRAAEYSMNVGFLGKGNSSKKQLVEQVAGATGFKLHEDWGTTPSIDCLSVADYD 241
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 183 LRAAEYSMNVLGFLAGNASNDASLADQIEAGATGFKLHEDWGTTPSIDCLSVADYD 242
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 242 VQVCIHDTVNEAGVDDTLNMGNGRAHYHIEGAGGCHGSPDVTIMAGELNIPSSSTTP 301
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 243 VQVAIHDTVNEAGVDDTLNMGNGRAHYHIEGAGGCHGSPDVTIMAGELNIPSSSTTP 302
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 302 TIPTVINTVAEHLDMLTCHHLDKRIREDLQFSOSRIRPGSIAAEDVLHDMGVIAMTSSD 361
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 303 TIPTVINTVAEHLDMLTCHHLDKRIREDLQFSOSRIRPGSIAAEDVLHDMGVIAMTSSD 362
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 362 SOAMGRAGEVTPRTWOTADKXKKEFGKLPEDGKNDNFRIKRYISKYTIINPAITHGVSEY 421
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 363 SOAMGRAGEVTPRTWOTADKXKKEFGKLPEDGKNDNFRIKRYISKYTIINPAITHGVSEY 422
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 422 IGSVEEGKIADLVNPAFFGKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFGHH 481
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 423 VGSVEGKADLVNPAFFGKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFGHH 482
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 482 GKAKFDTISITFVSKVAVENGVEKELGLERQVLVKNCRNITKQDFKNDTKAKIIVDPKT 541
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 483 GKAKYDANITFVSOAAVDKIKKEELGLERQVLVKNCRNITKQDFKNDTKAKIIVDPKT 542
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 542 FEVFDGKLTCKTSPQVPLAQRTFF 568
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 543 YHVPVFGKEVTSKPAKVSLAQLFSIF 569
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
```

RESULT 5

```
ID  _URE2_HELPY STANDARD;          PRT;  569 AA.
AC  P69996; P14917; Q9R3B3;
DT  04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT  04-JAN-2005, sequence version 1.
DT  07-MAR-2006, entry version 11.
DE  Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN  Name=ureB; Synonyms=hpuB; OrderedLocusNames=HP0072;
OS  Helicobacter pylori (Campylobacter pylori).
OC  Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC  Helicobacteraceae; Helicobacter.
OX  NCBI_TaxID=210;
RN  NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RP  STRAIN=85P;
RC  MEDLINE=91161505; PubMed=2001995;
RX  Labigne A., Cussac V., Courcoux P.;
RT  "Shuttle cloning and nucleotide sequences of Helicobacter pylori genes
   responsible for urease activity."
RL  J. Bacteriol. 173:1920-1931(1991).
RN  [2]
RP  NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC  STRAIN=CPM630;
RX  MEDLINE=90221820; PubMed=2326167;
RA  Clayton C.L., Pallen M.J., Kleanthous H., Wren B.W., Tabagchali S.;
RT  "Nucleotide sequence of two genes from Helicobacter pylori encoding
```

```
for urease subunits.";
RL  Nucleic Acids Res. 19:362-362(1990).
RN  [3]
RP  NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC  STRAIN=26695 / ATCC 700392;
RX  MEDLINE=97394467; PubMed=9252185; DOI=10.1038/41483;
RA  Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
   Fleischnann R.D., Ketchum K.A., Klenk H.-P., Gill S.R.,
   Dougherty B.A., Nelson K.E., Quackenbush J., Zhou L., Kirkness E.P.,
   Peterson S.N., Loftus B.J., Richardson D.L., Dodson R.J., Khalak H.G.,
   Glodek A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D.,
   Hickey B.K., Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D.,
   Kelley J.M., Cotton M.D., Weidman J.F., Fujii C., Bowman C.,
   Watthey L., Wallin E., Hayes W.S., Borodovsky M., Karp P.D.,
   Smith H.O., Fraser C.M., Venter J.C.;
RT  "The complete genome sequence of the gastric pathogen Helicobacter
   pylori.";
RL  Nature 388:539-547(1997).
RN  [4]
RP  PROTEIN SEQUENCE OF 1-30, AND KINETIC PARAMETERS.
RX  PubMed=1857197; DOI=10.1016/0882-4010(91)90062-F;
RA  Evans D.J. Jr., Evans D.G., Kirkpatrick S.S., Graham D.Y.;
RT  "Characterization of the Helicobacter pylori urease and purification
   of its subunits.";
RL  Microb. Pathog. 10:15-26(1991).
RN  [5]
RP  PROTEIN SEQUENCE OF 1-20.
RX  MEDLINE=90264448; PubMed=2188975;
RA  Dunn B.E., Campbell G.P., Perez-Perez G.I., Blaser M.J.;
RT  "Purification and characterization of urease from Helicobacter
   pylori.";
RL  J. Biol. Chem. 265:9464-9469(1990).
RN  [6]
RP  PROTEIN SEQUENCE OF 1-15.
RX  MEDLINE=90202165; PubMed=2318539;
RA  Hu L.T., Mobley H.L.F.;
RT  "Purification and N-terminal analysis of urease from Helicobacter
   pylori.";
RL  Infect. Immun. 58:992-998(1990).
RN  [7]
RP  PROTEIN SEQUENCE OF 1-12.
RC  STRAIN=NCCT 11637;
RX  PubMed=1452359;
RA  Turbett G.R., Hoej P.B., Horne R., Mee B.J.;
RT  "Purification and characterization of the urease enzymes of
   Helicobacter species from humans and animals.";
RL  Infect. Immun. 60:5259-5266(1992).
RN  [8]
RP  NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 565-569.
RC  STRAIN=85P;
RX  MEDLINE=92210488; PubMed=1313413;
RA  Cussac V., Ferrero R.L., Labigne A.;
RT  "Expression of Helicobacter pylori urease genes in Escherichia coli
   grown under nitrogen-limiting conditions.";
RL  J. Bacteriol. 174:2466-2473(1992).
RN  [9]
RP  X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS), AND SUBUNIT STRUCTURE.
RX  MEDLINE=21267477; PubMed=11373617; DOI=10.1038/88563;
RA  Ha N.-C., Oh S.-T., Sung J.Y., Cha K.A., Lee M.H., Oh B.-H.;
RT  "Supramolecular assembly and acid resistance of Helicobacter pylori
   urease.";
RL  Nat. Struct. Biol. 8:505-509(2001).
CC  -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC  -!- COFACTOR: Binds 2 nickel ions per subunit.
CC  -!- BIOPHYSICOCHEMICAL PROPERTIES:
   Kinetic parameters:
   KM=0.48 mM for urea;
   pH dependence:
   Optimum pH is 8.0. Active from pH 4.0 to 10.0;
CC  -!- SUBUNIT: Heterohexamer of 3 alpha and 3 beta subunits. Four
   heterohexamers assemble to form a 16 nm spherical complex.
CC  -!- PTM: Lys-219 is carbamylated. The carbamoyl group provides the
   ligands for the two nickel ions (By similarity).
```



```

RESULT 6
ID Q950Q5 HELPY PRELIMINARY; PRT; 569 AA.
AC Q950Q5
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Urease B.
GN Name=ureB;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=HPK5;
RC MEDLINE=20305050; PubMed=10844692;
RA Akada J.K., Shirai M., Takeuchi H., Tsuda M., Nakazawa T.;
RT "Identification of the urease operon in Helicobacter pylori and its
RT control by mRNA decay in response to pH.";
RL Mol. Microbiol. 36:1071-1084(2000).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC EMBL; AB032429; BAA84533.1; -; Genomic_DNA.
DR HSSP; P14917; 189Y.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0019627; P:urea metabolism; IEA.
DR InterPro; IPR006680; Amidohydro 1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR011612; Urease_alpha N.
DR InterPro; IPR008295; Urease_alpha1.
DR Pfam; PF01979; Amidohydro 1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF01226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alpha; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
SQ SEQUENCE 569 AA; 61671 MW; 4C8A6BDA3A295584 CRC64;

Query Match 76.6%; Score 2297; DB 2; Length 569;
Best Local Similarity 73.4%; Pred. No. 1.6e-42;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

Qy 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVHDYTYGELKFGAGKTIREGMGOSNSPD 61
Db 3 KISRKEYSMYPTTGDVKRLGDTDLIAEVEHDYTYGELKFGGKTIREGMSQSNPS 62
Qy 62 ENTLDLVITNAMIIDYTGIIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVMVGVEALA 121
Db 63 KEELDLIITNALIVDYGIIYKADIGIKGKIAGIKGGNKMDDGVKNLSVGPATEALA 122
Qy 122 GEGMIITAGGIDSHTHFLSPQPFPTALANGVTTMTFGGGTGPVDGNTATTTPGKNLHRM 181
Db 123 GEGLIIVTAGGIDTHIHFIHFISPOQIPTAFASGVTTMTGGGTGPADGNTATTTPGRNLKWM 182
Qy 182 LRAAEYSNMVGFGLCKGNSSSKKQLVEVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEYD 241
Db 183 LRAAEYSNMNIGFLAKGNASNDASLADQIEAGAIKFKIHEWDGTTTPSAINHALDVADKYD 242
Qy 242 VQVCIIHTDVNEAGYVDDTLNAMNGRAIHAYHIEGAGGCHSPDVTITMAGELNIPSSSTTP 301
Db 243 VQVAIHTDLNEAGCVETMAAIGRTWHTFTHEGAGGCHAPDIIKVAGEHNILPASTNP 302
Qy 302 TIPTYTINTVAEHLDMMLTCHLLDKRIREDLPQSQSRIRPGSIAAEDVLHDMGVIAWTSSD 361
Db 303 TIPTVTNTEAEHMDMLVCHLLDKSIKEDVQPADSRIRPQTIAAEDTLHDMGIFSISSD 362

Qy 362 SQAMGRAGEVIRPTWOTADKNKKFKLPEDEKNDNDFRIKRYISKYINPALTHTGVSEY 421
Db 363 SQAMGRVGEVIRPTWOTADKNKKFKLPEDEKNDNDFRIKRYISKYINPALTHTGVSEY 422
Qy 422 TGSVEEGKIADLVNPPAFGKPKIVIKGMVVMSEMGDSNASVPTPQPVYVREMGHH 481
Db 423 VGSVEVGKVDLVLWSPAFFGVKPNMIKGGFIALSQMGDANASITPQPVYVREMFHH 482
Qy 482 GKAKEDTSTTFVSKVAYENGKVKGLERQVLVVKNCRNITKKDFKNDKTKAKITVDPKT 541
Db 483 GKAKYDANITTFVSOAYDKIKERGLERQVLVVKNCRNITKKDMQFNDDTAHEVNPET 542
Qy 542 FEVFDGKLCSTKSTPSQVPLAQRYTF 568
Db 543 YHVFVDGKEVTSKPKATKVSALQFSIF 569

RESULT 7
Q38Q64 HELPY PRELIMINARY; PRT; 568 AA.
AC Q38Q64;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Urease B (Fragment).
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=NCTC 11639;
RA Li Y., Ning Y.-S., Long M., Wang Y.-D., Dong W.-Q., Li M.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC EMBL; DQ141576; ABA87133.1; -; Genomic_DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0019627; P:urea metabolism; IEA.
FT NON_TER 1
FT NON_TER 568
SQ SEQUENCE 568 AA; 61524 MW; 5CD9DC062D41P883 CRC64;

Query Match 76.5%; Score 2293; DB 2; Length 568;
Best Local Similarity 73.2%; Pred. No. 2e-42;
Matches 415; Conservative 76; Mismatches 76; Indels 0; Gaps 0;

Qy 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVHDYTYGELKFGAGKTIREGMGOSNSPD 61
Db 2 KISRKEYSMYPTTGDVKRLGDTDLIAEVEHDYTYGELKFGGKTIREGMSQSNPS 61
Qy 62 ENTLDLVITNAMIIDYTGIIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVMVGVEALA 121
Db 62 KEELDLIITNALIVDYGIIYKADIGIKGKIAGIKGGNKMDDGVKNLSVGPATEALA 121
Qy 122 GEGMIITAGGIDSHTHFLSPQPFPTALANGVTTMTFGGGTGPVDGNTATTTPGKNLHRM 181
Db 122 GEGLIIVTAGGIDTHIHFIHFISPOQIPTAFASGVTTMTGGGTGPADGNTATTTPGRNLKWM 181
Qy 182 LRAAEYSNMVGFGLCKGNSSSKKQLVEVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEYD 241
Db 182 LRAAEYSNMNIGFLAKGNASNDASLADQIEAGAIKFKIHEWDGTTTPSAINHALDVADKYD 241
Qy 242 VQVCIIHTDVNEAGYVDDTLNAMNGRAIHAYHIEGAGGCHSPDVTITMAGELNIPSSSTTP 301
Db 242 VQVAIHTDLNEAGCVETMAAIGRTWHTFTHEGAGGCHAPDIIKVAGEHNILPASTNP 301
Qy 302 TIPTYTINTVAEHLDMMLTCHLLDKRIREDLPQSQSRIRPGSIAAEDVLHDMGVIAWTSSD 361

```


Db	363	SOAMRGVGEVITRTWQADKNKCFGRLEKEEGKGNNDNRIKRYLSKYTINPAIAHGISEY	422
Qy	422	IGSVEEGKIALDLVWNPAFFGVGPKVIKGGMVVFSEMGDSNASVPTQPQVYVYREMPFGHH	481
Db	423	VGSVEVGKVALDLVWSPAFFGVGKFNMIKGGFIALSQMGDNANASIPTPQPYVYREMPFAHH	482
Qy	482	GKAKFDTSITFVSKVAYENGVEKGLGLERQVLPVKNCRNITKKDPKFNKD TAKITVDPKT	541
Db	483	GKAKYDANITFVSQAAYDKGIKEBELGLERQVLPVKNCRNITKKDMQFNDTTAHIEVNPE	542
Qy	542	FEYFVDGKLCIT 552	
Db	543	YHVFVDGKEVT 553	

RESULT_12

URE2_HELPFE

ID

URE2_HELPFE

STANDARD;

PRT;

569 AA.

AC

Q08716; Q093B3;

DT

01-OCT-1994, integrated into UniProtKB/Swiss-Prot.

DT

01-OCT-1994, sequence version 1.

DT

07-FEB-2006, entry version 47.

DE

Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).

GN

Name=ureB;

OS

Helicobacter felis.

OC

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

OC

Helicobacteraceae; Helicobacter.

OX

NCBI_TaxID=214;

RN

[1]

RP

NUCLEOTIDE SEQUENCE [GENOMIC DNA].

RP

MEDLINE=94018627; PubMed=8412683;

RA

Ferrero R.L., Labigne A.;

RT

"Cloning, expression and sequencing of Helicobacter felis urease

RT

genes.";

RL

Mol. Microbiol. 9:323-333(1993).

RL

[2]

RP

PROTEIN SEQUENCE OF 1-12.

RP

STRAIN=ATCC 49179;

RC

PubMed=1452359;

RX

PubMed=1452359;

RA

Turbett G.R., Hoef P.B., Horne R., Mee B.J.;

RT

"Purification and characterization of the urease enzymes of

RT

Helicobacter species from humans and animals.";

RL

Infect. Immun. 60:5259-5266(1992).

CC

-1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).

CC

-1- COFACTOR: Binds 2 nickel ions per subunit (By similarity).

CC

-1- SUBUNIT: Consists of two subunits (alpha and beta).

CC

-1- PTM: Lys-219 is carbamylated. the carbamoyl group provides the

CC

ligands for the two nickel ions (By similarity).

CC

-1- SIMILARITY: Belongs to the urease family.

CC

-1- CAUTION: In Helicobacter the beta subunit is what is known in

CC

other bacteria as the alpha subunit.

CC

CC

Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC

Distributed under the Creative Commons Attribution-NoDerivs License

CC

EMBL;

X69080; CAA48826.1; -; Genomic_DNA.

DR

PIR; C49215; C49215.

DR

FIR; S35291; S35291.

DR

HSSP; P14917; 1B9Z.

DR

InterPro; IPR006680; Amidohydro_1.

DR

InterPro; IPR005848; Urease_alpha.

DR

InterPro; IPR011612; Urease_alpha_N.

DR

InterPro; IPR008295; Urease_alpha1one.

DR

Pfam; P001979; Amidohydro_1; 1.

DR

Pfam; P000449; Urease_alpha; 1.

DR

PIRS; PIRS001226; Urease_alpha; 1.

DR

PRINTS; PR01752; UREASE.

DR

TIGRFAMs; TIGR01792; urease_alpha; 1.

DR

PROSITE; PS01120; UREASE_1; 1.

DR

PROSITE; PS00145; UREASE_2; 1.

DR

Direct protein sequencing; Hydrolase; Metal-binding; Nickel.

FT

CHAIN 1 569 Urease beta subunit.

FT				/FTID=PRO_0000067530.
FT	ACT SITE	322	322	Proton donor (By similarity).
FT	METAL	136	136	Nickel 2 (By similarity).
FT	METAL	138	138	Nickel 2 (By similarity).
FT	METAL	219	219	Nickel 1 (By similarity).
FT	METAL	219	219	Nickel 2 (By similarity).
FT	METAL	248	248	Nickel 1 (By similarity).
FT	METAL	274	274	Nickel 1 (By similarity).
FT	METAL	362	362	Nickel 2 (By similarity).
FT	BINDING	221	221	Substrate (By similarity).
SQ	SEQUENCE	569 AA;	61703 MW; 66DBB99E2CC7CER3 CRC64;	
	Query Match	74.8%;	Score 2244;	DB 1; Length 569;
	Best Local Similarity	72.5%;	Pred. No. 2.5e-41;	
	Matches 411;	Conservative 71;	Mismatches 85;	Indels 0; Gaps 0;
QY	2 KMKQEVNTYPTKGDVRLGDTDLWAEVEHDYTYYGEELKEGAGKTIREGMQSHPD	61		
DB	3 KISRKEYVMYGPTTGDVRVLGDTDLLEVEHDCITTYGEEIKFGGGKTIRDMSQTSNPS	62		
QY	62 ENTLDELIVTNAMIIDVTGIYKADIGIKNGKIHGIKRAGNKMDQGVSPHMVVGVGTTEALA	121		
DB	63 SYELDLVLINALIVDTGIYKADIGIKOGKIAGIKRAGNKMDQGVDDNNLVCVPATEALA	122		
QY	122 GEGMIITAGIDSHTFLSPQQPPTALANGVTMTMFGGTPGVGTNATTIPGKNILHRM	181		
DB	123 AGLIIVTAGIDITHIFISPPQIPTAPASGVTTMIGGCTPADGTNATTITPGRANLKSM	182		
QY	182 LRAAEYSNMVFLGKGNSSKKQLVEOVFAGAIGFKLEDWGTTTPSAIDHCLSLVADEYD	241		
DB	183 LRAAEYAMNLGFLAKGNVSYPESLRDQIEBAGAIGFKIHEDWGSTPAAIHHCLNLVADEYD	242		
QY	242 VQVCIIHTDVNEAGYVDDTLNAMNGRAIHAYHIETEGAGCGHSPOVITWAGELNILPSSTTP	301		
DB	243 VQVALHTDLNEAGCVEDTLEA TAGRTIHTFHETEGAGCGHAPDVIKWAGEFNILPASTNP	302		
QY	302 TIPYTIINTVAEHLDMLTCHHLDKRINREDLOFSOSRIIRPGSIAAADVLHDMGVIAMTSSD	361		
DB	303 TIFTKNTAEHMDMLMVCHLDKS IKEDYQFADSRIRPQTAAEDQLHDMGIFSITSSD	362		
QY	362 SQAMGRAGEV IPRWTQTADNKKVEFGKLPRDGKDNDNFRIKRYISKYTIINPALTHGVSEY	421		
DB	363 SQAMGRVGVEITWTQTADNKKVEFGKLSEKGDNDFRIKRYISKYTIINPGIAHGISDY	422		
QY	422 IGSVBEKGIADLVVNPNPAPPFGVKPKV IKGWVVFSBMGSDNASVPTQPQVYYREMFGHH	481		
DB	423 VGSVEVGKYADLVNWSPAFFGIKPMMIKGFALTALSOMGDANASIPITQPQVYYREMFGHH	482		
QY	482 GKAFDPTSIITFVSKVAYENGVEKGLGERQVL PVKNCRNITTKDKFKENDKTAKITVDPKT	541		
DB	483 GKMKFDNIITFVSOQAYKAGIKEELGLDRAAPPVKNCRNITTKDKLFENDVTAHIDVNPET	542		
QY	542 FEVFVQGLCTSKPTSOVPLAQRYTFP	568		
DB	543 YKVQVDGKEVTSKAADLSLAOLYNLF	569		
	RESULT 13			
	Q93PJ4_HELHP			
ID	Q93PJ4_HELHP	PRELIMINARY;	PRT;	569 AA.
AC	Q93PJ4; Q7BYA2;			
DT	O1-DSC-2001,	integrated into UniProtKB/TREMBL.		
DT	O1-DSC-2001,	sequence version 1.		
DT	O7-FRB-2006,	entry version 21.		
DE	Urease beta subunit ureB (EC 3.5.1.15).			
GN	Names:ureB; Ordered locus names=HH0408; ORFNames=HH_0408;			
OS	Helicobacter hepaticus.			
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;			
OC	Helicobacteraceae; Helicobacter.			
OX	NCBI_TaxID=32025;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RX	DEFAULTLINE=21391860; PubMed=11500473;			

DOI=10.1128/IAI.69.9.5914-5920.2001;	
Beckwith C.S., McGee D.J., Mobley H.L.T., Riley L.K.;	
"Cloning, expression, and catalytic activity of Helicobacter hepaticus	
urease".	
RL Infect. Immun. 69:5914-5920(2001).	
RL [2]	
RL NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].	
RC STRAIN=ATCC 51449 / 3B1;	
RC MEDLINE=2709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;	
RA Surbaum S., Josenhans C., Stenzenbach T., Drescher B., Brandt P.,	
RA Ball M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,	
RA Holland R., Klein K., Koenig J., Macko U., Mendz G.L., Nyakatura G.,	
RA Schauer D.B., Shen Z., Weber J., Froesch M., Fox J.G.;	
RT "The complete genome sequence of the carcinogenic bacterium	
RT Helicobacter hepaticus.";	
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).	
CC -----	
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms	
CC Distributed under the Creative Commons Attribution-NonDerivs License	
CC -----	
CC EMBL; AF312656; AAK69199.1; -; Genomic DNA.	
DR EMBL; AE017125; AAP77005.1; -; Genomic DNA.	
DR HSSP; P14917; I8Y9.	
DR HSSP; P18314; IASN.	
DR BioCyc; HHEP235279:HH0408-MONOMER; -.	
DR GO; GO:0016787; F:hydrolase activity; IEA.	
DR GO; GO:0016151; F:nickel ion binding; IEA.	
DR GO; GO:0009039; F:urease activity; IEA.	
DR GO; GO:0019627; P:urea metabolism; IEA.	
DR InterPro; IPR006690; Amidohydro 1.	
DR InterPro; IPR005848; Urease_alpha N.	
DR InterPro; IPR011612; Urease_alpha N.	
DR InterPro; IPR008295; Urease_alpha1one.	
DR Pfam; PF001979; Amidohydro 1; 1.	
DR Pfam; PF00449; Urease_alpha; 1.	
DR Pfam; PF00449; Urease_alpha; 1.	
DR Pfam; PF00449; Urease_alpha; 1.	
DR PRINTS; PRS0001226; Urease_alpha; 1.	
DR PRINTS; PRS01752; UREASE.	
DR TIGRFAMs; TIGR01792; urease_alph; 1.	
DR PROSITE; PS01120; UREASE_1; 1.	
DR PROSITE; PS00145; UREASE_2; 1.	
DR PROSITE; PS00145; UREASE_2; 1.	
DR Complete proteome; HydroLase.	
SW SEQUENCE 569 AA; 61510 MW; 2825DLE6406F6429 CRC64;	
SW -----	
Query Match	70.3%; Score 2109; DB 2; Length 569;
Best Local Similarity	68.1%; Pred. No. 2.8e-38;
Matches 388; Conservative	83; Mismatches 95; Indels 4; Gaps 6;
Qy 1	MKKKKQBYVNTYGTGDKVRLGDTDLMAEVEHYDTTYGEEKLFGAGKTIRMGQSNRP 60
Db 2	IKISRKQYASMYGPTGDKVRLGDTNLPFAETEKDYTLGEEIKFGGKTI RDGMAQAS 61
Qy 61	DENTLDVATNMIIDYTGIVKADIGIKNGKHIGKAGNKMDOGSPHVVGVCTEAL 120
Db 62	YTMELDVAITNMIIDYTGIVKADIGIKGKIVGKAGNPDQDSVNEAMVVGAATEVI 121
Qy 121	AGEGMIITAGGIDSHTHFLSPQFPPTALANGVTWTFGGGTGPVDCGTNATTTIPGKWNLR 180
Db 122	AGEQIITAGGIDTHIHFIPTQIPTALYSGVTTWIGGTPAAGTNAATCTPGKWNHQ 181
Qy 181	MLRAEYYSNMVGLFGKNGSSKKOLVQVEBAGAIQFKLHEDWGTTPSAIDHCLSVADYE 240
Db 182	MLRAAESYAMNLGFFGKNGSNSEBLEBQIKAGALGLKVHEDWGSTPAINHALNVAQKY 241
Qy 241	DVQVCITHTDVTNEAGYVDDTLNANMGRAIHAYHIEGAGGSHSPDVIWAGELNII 300
Db 242	DVQVAIHTDITLNEAGCVDEWTWKAIDGRTHFTFHTEGAGGGHAPDIKAAGEPNII 301
Qy 301	PTPTPTNTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEVLHDMGVIAWTS 360
Db 302	PTPTPTNTADEHLDMVMCHLDKIKEDVAFADSRIRPETIAAEDTLHDMGIFSI 361
Qy 361	DSQAMGRAGEVIPRTWQTADKQKKEFGKLPEB-D-GKNDNFRIKRYISKYITINPALTHGVS 419

Db 362 DSQAMRGVGEVITRTWQTADCKNEFGALKKEBCG-ENDNFRIKRYISKYTINPAIAHGIS 420

Qy 420 EYIGSVBECKIADLVVWNPAPFGVCPKIVKGMVVFSEMGDSNASVTPQPVYREMF 479

Db 421 EYIGSVBECKIADLVVWNPAPFGVCPKIVKGMVVFSEMGDSNASVTPQPVYREMF 480

Qy 480 HHGKAKPDTSTFVSKVAYENGVEKGLERQVLVVKNCRNITKDKPFNDKTAKITVDP 539

Db 481 SYGKAKNCALTFVSKIAYDCHKEBELGLERILLVVKNCRNITKDKPFNDVITPIEVNP 540

Qy 540 KTFEYFVDG-KLCTSKTSQVPLAQRYTF 568

Db 541 EYIEVRVNNTKI-TSKPVEKVSGLQYLCIF 569

RESULT 14

URE1_BACSB STANDARD; PRT; 569 AA.

AC Q07397; 01-OCT-1994, integrated into UniProtKB/Swiss-Prot.

DT 01-OCT-1994, sequence version 1.

DT 07-FEB-2006, entry version 51.

GN Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase).

GN Name-ureC;

OS Bacillus sp. (strain TB-90).

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=36824;

RN [1]

RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND PARTIAL PROTEIN SEQUENCE.

RX MEDLINE=94117379; PubMed=8288539; Masaki H., Uozumi T.; Maeda M., Hidaka M., Nakamura A., Masaki H., Uozumi T.; "Cloning, sequencing, and expression of thermophilic Bacillus sp. strain TB-90 urease gene complex in Escherichia coli."; J. Bacteriol. 176:432-442(1994).

RL J. Bacteriol. 176:432-442(1994).

CC -1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).

CC -1- COFACTOR: Binds 2 nickel ions per subunit (By similarity).

CC -1- SUBUNIT: (Alpha, beta, gamma) (3) (By similarity).

CC -1- PFM: Lys-219 is carbamylated. The carbamoyl group provides the ligands for the two nickel ions (By similarity).

CC -1- SIMILARITY: Belongs to the urease family.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

CC EMBL: D14439; BAA03325.1; -; Genomic_DNA.

DR PIR: C36950; C36950.

DR HSP: P18314; 1EJX.

DR InterPro: IPR006680; Amidohydro 1.

DR InterPro: IPR011550; Amidohydro like.

DR InterPro: IPR005848; Urease_alpha.

DR InterPro: IPR011612; Urease_alpha N.

DR InterPro: IPR008295; Urease_alpha1one.

DR Pfam: PF01979; Amidohydro_1; 1.

DR Pfam: PF00449; Urease_alpha; 1.

DR PRSF: PRSF001226; Urease_alpha; 1.

DR PRINTS: PR01752; UREASE.

DR ProDom: PD001248; Amidohydro like; 1.

DR TIGRfams: TIGR01792; urease_alpha; 1.

DR PROSITE: PS01120; UREASE_1; 1.

DR PROSITE: PS00145; UREASE_2; 1.

KW Direct protein sequencing; Hydrolase; Metal-binding; Nickel.

FT CHAIN 1 569 Urease alpha subunit.

FT FTID=PRO_000067537.

FT ACT_SITE 322 322 Proton donor (By similarity).

FT METAL 136 136 Nickel 2 (By similarity).

FT METAL 138 138 Nickel 2 (By similarity).

FT METAL 219 219 Nickel 1 (By similarity).

FT METAL 219 219 Nickel 1 (By similarity).

FT METAL 248 248 Nickel 1 (By similarity).

FT METAL 274 274 Nickel 1 (By similarity).

FT METAL 362 362 Nickel 2 (By similarity).

FT BINDING 221 221 Substrate (By similarity).

SQ SEQUENCE 569 AA; 61397 MW; 38AA83133A31A49D CRC64;

Query Match 68.2%; Score 2044; DB 1; Length 569;

Best Local Similarity 66.4%; Pred. No. 8.1e-37;

Matches 378; Conservative 76; Mismatches 108; Indels 7; Gaps 4;

Qy 3 MKKQEVNTYVPTGPKDKVGLDGLMAEVEHVDVTTTGEELKFGAGKTIREGMQ---SNS 59

Db 5 MRKQVADMFPTVGDALRLADSELFIETIEKDYTTVGDEVKFGGKVIKRGMGHPLATS 64

Qy 60 PDENTLDLVITNAMIIDYTIYKADIGIKNGKIHGIGKAGNKMQDQGVSPHVMVGVGTEA 119

Db 65 -DE-CVDLVLTNAIIVDYTIYKADIGIKDGMIASIGKAGNPLLMQGV--MVIGAAATEV 120

Qy 120 LAGEGMIIITAGGIDSHFLSPQOFTALANGVTTMFGGGTGPVVDGNTATITPKGNLH 179

Db 121 IAAEGMIVTAGGIDAHIFICPOQIETALASGVTTMIGGGTGPATGTNATTCTGPNIIH 180

Qy 180 RMLRAAEYSVMVGLKGNSSKKOLVQVEAGAIGFKLHEDWGTTPPSAIDHCLSVAD 239

Db 181 RMLQAAEFPINLGLKGNCSDEAPLKEIQEAGAVGLKLUHEDWGSTAAAIIDTCLKVADR 240

Qy 240 YDVQCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVITMAGELNLPSS 299

Db 241 YDVQVAILHTDLNEGGFVEDTLKIDGRVHTVHTGAGGGGHAPDIIKAAGFPNLPSS 300

Qy 300 TPTPIYTTINVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAABDVLDHMGVIAMTS 359

Db 301 NTPRPYTTINLEHLDMLMVCCHLDANIPEDIAFADSRIRKETIAABDVLDHDLGVFSMIS 360

Qy 360 SDSQAMGRAGEVIPRTWQTADCKNEFGALKKEBCGKNDNFRIKRYISKYTINPAIAHGVS 419

Db 361 SDSQAMGRVGEVITRTWQTADCKNEFGALKKEBCGKNDNFRIKRYISKYTINPAIAHGIA 420

Qy 420 EYIGSVBECKIADLVVWNPAPFGVCPKIVKGMVVFSEMGDSNASVTPQPVYREMF 479

Db 421 DVGSEVEGKLADLVVWNPAPFGVCPKIVKGMVVFSEMGDSNASVTPQPVYREMF 480

Qy 480 HHGKAKPDTSTFVSKVAYENGVEKGLERQVLVVKNCRNITKDKPFNDKTAKITVDP 539

Db 481 AKGDAKYQTSITFVSKAAVEKHEQLGKVKPVHGIRKLTOKDLILNDKPKIDVDP 540

Qy 540 KTFEYFVDGKLCTSKTSQVPLAQRYTF 568

Db 541 QTYEVKVDQLVTCBPABIVPMAQRYFLF 569

RESULT 15

Q9KG59_BACHD PRELIMINARY; PRT; 571 AA.

AC Q9KG59; 01-OCT-2000, integrated into UniProtKB/TrEMBL.

DT 01-OCT-2000, sequence version 1.

DT 07-FEB-2006, entry version 25.

DE Urease alpha subunit (EC 3.5.1.5).

GN Name-ureC; OrderedLocusNames=BH0254;

OS Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=86665;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=C-125 / JCM 9153;

RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N., Fuji F., Hirano C., Nakamura Y., Ogasawara N., Kuhara S., Horikoshi K.;

RA "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis.";

RL Nucleic Acids Res. 28:4317-4331(2000).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

CC EMBL: BA000004; BAB03973.1; -; Genomic_DNA.

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:30:32 ; Search time 194 Seconds
(without alignments)
615.121 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVQLVQSGAEVKPGASV.....KVIKRGGLGLVYKDDDDK 261

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 2589679 seqs, 457216429 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2589342

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_8.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*
- 10: Geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	198	75.9	260	9 AEC92007	Aec92007 Single ch
2	51	19.5	110	8 ADW24449	Adw24449 HCMV-Fab-
3	48	18.4	93	8 ADU04216	Adu04216 Human imm
4	48	18.4	100	6 AAE35907	Aae35907 Human Al7
5	48	18.4	100	6 ABO27142	Abo27142 Human ger
6	48	18.4	100	7 ADF10189	Adf10189 Antibody
7	48	18.4	100	7 ADF10087	Adf10087 VEGF anti
8	48	18.4	100	7 ADF09982	Adf09982 Antibody
9	48	18.4	100	7 ADJ80262	Adj80262 V kappa ge
10	48	18.4	100	8 ADO07331	Ado07331 Human ant
11	48	18.4	100	9 ADY75436	Ady75436 Human ger
12	48	18.4	100	9 AEB13589	Aeb13589 Human var
13	48	18.4	100	9 AEB13642	Aeb13642 Human var
14	48	18.4	100	9 AED04308	Aed04308 Human ant
15	48	18.4	100	9 AED06937	Aed06937 Human ant
16	48	18.4	100	10 AEE47598	Aee47598 Human CDR
17	48	18.4	100	10 AEF18501	Aef18501 HAHa prot
18	48	18.4	111	4 AAE06948	Aae06948 Human HF-
19	48	18.4	111	4 AAE06996	Aae06996 Human kap
20	48	18.4	111	4 AAU09920	Aau09920 Human ant
21	48	18.4	111	7 ABR61476	Adr61476 Human ant
22	48	18.4	111	8 ADQ89233	Adq89233 Human imm
23	48	18.4	111	8 ADQ89281	Adq89281 Human imm

24	48	18.4	111	9 ADU26553	Adu26553 Human ant
25	48	18.4	111	9 AEB09554	Aeb09554 Human lig
26	48	18.4	111	9 AEB09506	Aeb09506 Human HF-
27	48	18.4	111	9 AEC92164	Aec92164 Human HF-
28	48	18.4	112	2 AAW79227	Aaw79227 Light cha
29	48	18.4	112	2 AAW53585	Aaw53585 Light cha
30	48	18.4	112	4 AAE07003	Aae07003 Human kap
31	48	18.4	112	4 AAE06993	Aae06993 Human kap
32	48	18.4	112	4 AAE06999	Aae06999 Human kap
33	48	18.4	112	5 ADF98232	Adf98232 Human HF-
34	48	18.4	112	8 ADQ89284	Adq89284 Human imm
35	48	18.4	112	8 ADQ89278	Adq89278 Human imm
36	48	18.4	112	8 ADQ89288	Adq89288 Human imm
37	48	18.4	112	9 ADX01284	Adx01284 Human Hu5
38	48	18.4	112	9 AEB09557	Aeb09557 Human lig
39	48	18.4	112	9 AEB09551	Aeb09551 Human lig
40	48	18.4	112	9 AEB09561	Aeb09561 Human lig
41	48	18.4	113	8 ADO07283	Ado07283 Human pro
42	48	18.4	113	9 ADZ42043	Adz42043 Ig L chai
43	48	18.4	113	9 ADZ42042	Adz42042 Ig L chai
44	48	18.4	113	9 ADZ42041	Adz42041 Ig L chai
45	48	18.4	121	10 AEG01367	Aeg01367 Kallikrei

ALIGNMENTS

RESULT 1
AEC92007
ID AEC92007 standard; protein; 260 AA.
XX
AC AEC92007;
XX
DT 01-DEC-2005 (first entry)
XX
DE Single chain variable fragment antibody BC-12 protein SEQ ID NO 2.
XX

cytostatic; immunostimulant; pharmaceutical; protein interaction;
therapeutic; immune modulation; solid tumor; leukemia; colorectal
tumor; breast tumor; uterine cervix tumor; uterine fibroids; ovary tumor;
gynecology and obstetrics; polycystic ovary syndrome; antiinfertility;
gynecological; polyp; growth disorder; prostate tumor; prostate disease;
androgyny; genitourinary disease; pituitary tumor; endocrine disease;
KW adenocarcinoma; melanoma; bone tumor; musculoskeletal disease;
KW multiple myeloma; hematological disease; immune disorder;
KW central nervous system tumor; neurological disease; glioma;
KW astroblastoma; neoplasm; single chain antibody; BC-12.

OS Homo sapiens.
OS Synthetic.

Key Location/Qualifiers
FT Misc-difference 1..2 "Encoded by ATG"
FT Misc-difference 62
/note= "Encoded by GCA"

PN WO2005091805-A2.

XX 06-OCT-2005.

PD 14-FEB-2005; 2005WO-US004612.

XX 13-FEB-2004; 2004US-0544807P.

PR 10-NOV-2004; 2004US-0626726P.

PR 11-FEB-2005; 2005US-00056825.

XX (SCKI) SCRIPPS RES INST.

XX Felding-Habermann B, Janda KD, Saven A;

PI WPI; 2005-664872/68.

DR N-PSDB; AEC92006.

XX New scFv Bc-12 or Bc-15 antibody comprises ligand mimetic, which
PT specifically binds to activated alphavbeta3 integrin receptor, useful for
PT treating, e.g. neoplastic disease, solid tumor, leukemia, or benign or
PT malignant breast cancer.
XX
PS Claim 2; SEQ ID NO 2; 109pp; English.
XX
CC The invention describes an antibody comprising a ligand mimetic, which
CC specifically binds to an activated alphavbeta 3 integrin receptor
CC differentially produced on a cell in a metastatic state compared to a
CC similar, non-metastatic cell, and does not bind to a non-activated
CC alphavbeta 3 integrin receptor. Also described are: a pharmaceutical
CC composition comprising the antibody; treating a disease state in a mammal
CC; a cell line comprising a tumor cell variant with a metastatic homing
CC propensity to a target tissue; producing an antibody phage population
CC having affinity for a tumor cell target; detecting tumor cells in a
CC mammal by treatment with a cancer therapeutic; inducing or enhancing an
CC immune response to an antigen in a mammal; detecting an activated cell
CC surface receptor on a metastatic tumor cell surface in a mammalian tissue
CC sample; interfering with cells liable to undergo metastasis associated
CC with a disease state; identifying cells liable to undergo metastasis
CC associated with a disease state; an isolated Bc-12 or Bc-15
CC polynucleotide comprising a nucleotide sequence that has at least 90%
CC percent identity to a fully defined 802 or 752 bp sequence (SEQ ID NO. 1
CC or 3) given in the specification; an isolated polypeptide comprising a
CC nucleotide sequence that has at least 90% sequence identity to SEQ ID NO.
CC 1 or 3, or shares a biological function with Bc-12 or Bc-15; a vector
CC comprising the polynucleotide; an expression vector comprising the
CC polynucleotide in which the nucleotide sequence of the polynucleotide is
CC operatively linked with a regulatory sequence that controls expression of
CC the polynucleotide in a host cell; a host cell comprising the
CC polynucleotide or progeny of the cell; and determining anti-metastatic
CC activity of a test compound in a mammal. The antibody is useful for
CC treating neoplastic disease, solid tumor, hematological malignancy,
CC leukemia, colorectal cancer, benign or malignant breast cancer, uterine
CC cancer, uterine leiomyomas, ovarian cancer, endometrial cancer,
CC polycystic ovary syndrome, endometrial polyps, prostate cancer, prostatic
CC hypertrophy, pituitary cancer, adenomyosis, adenocarcinomas, meningioma,
CC melanoma, bone cancer, multiple myeloma, CNS cancer, glioma, or
CC astroblastoma. This is the amino acid sequence of single chain variable
CC fragment antibody Bc-12.
XX
SQ Sequence 260 AA;

Query Match 75.9%; Score 198; DB 9; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.8e-156; Indels 0; Gaps 0;
Matches 198; Conservative 0; Mismatches 0;
QY 62 QKFGQRTMTTDTSTAYMELRSLRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGGG 121
DB 63 QKFGQRTMTTDTSTAYMELRSLRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGGG 122
QY 122 GSGGGGGGGGSEIVLTQSPVLTSLPVLTLQSPASISCRSSQNLVYSGNTYLSWFOQRPQS 181
DB 123 GSGGGGGGGGSEIVLTQSPVLTSLPVLTLQSPASISCRSSQNLVYSGNTYLSWFOQRPQS 182
QY 182 PRLIYKVSNRDGVDPFRFSGSGGTDFTLKISRVEAEDIGVYCMQGTTHWPPRTFGQGT 241
DB 183 PRLIYKVSNRDGVDPFRFSGSGGTDFTLKISRVEAEDIGVYCMQGTTHWPPRTFGQGT 242
QY 242 KVEIKRGLGLGLVDYKDD 259
DB 243 KVEIKRGLGLGLVDYKDD 260

RESULT 2
ADW24449
ID ADW24449 standard; protein; 110 AA.
XX
AC ADW24449;
XX
DT 10-MAR-2005 (first entry)

XX HCMV-Fab-6 amino acid sequence #1.
DE Humanized antibody; antibody engineering; fab; cytomegalovirus infection.
XX
OS Human herpesvirus 5.
XX
PN CN1445243-A.
XX
PD 01-OCT-2003.
XX
PF 13-DEC-2002; 2002CN-00155426.
XX
PR 13-DEC-2002; 2002CN-00155426.
XX
PA (VIRU-) INST VIRUS PREVENTION & CONTROL CHINA DI.
XX
PI Liang M, Gu S, Duan T;
XX
XX WPI; 2004-091790/10.
DR N-PSDB; ADW24448.
XX
PT Fab antibody of human resourced neutrality genetic engineering for anti
human cytomegalo virus.
XX
PS Disclosure; Fig 3b; 15pp; Chinese.
XX
CC The invention relates to a humanized neutral genetically engineering Fab
antibody (HCMV-Fab-2 and HCMV-Fab-6) of human cytomegalovirus. The Fab
fragment, gene product and application are also disclosed. The
recombinant antibody is determined by the specific gene sequence in the
hypervariable region of the light-chain and heavy-chain variable regions.
The antibody can be used for preventing and treating the diseases
associated with HCMV infection. The current sequence represents a HCMV-
Fab-6 amino acid sequence.
XX
SQ Sequence 110 AA;
Query Match 19.5%; Score 51; DB 8; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.6e-34;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 YLSWFOQRPQSPRLIYKVSNRDGVDPFRFSGSGGTDFTLKISRVEAED 220
DB 35 YLSWFOQRPQSPRLIYKVSNRDGVDPFRFSGSGGTDFTLKISRVEAED 85
RESULT 3
ADU04216
ID ADU04216 standard; protein; 93 AA.
XX
AC ADU04216;
XX
DT 13-JAN-2005 (first entry)
XX
DE Human immunoglobulin kappa light chain protein 5.
XX
KW viral infection; virucide; autoimmune disease;
KW lymphoproliferative disorder; vaccine; gene therapy; immunoglobulin;
KW antibody.
XX
OS Homo sapiens.
XX
PN WO2004090544-A2.
XX
PD 21-OCT-2004.
XX
PF 13-APR-2004; 2004WO-CA000544.
XX
PR 09-APR-2003; 2003US-0461137P.
PR 30-SEP-2003; 2003US-0506779P.
XX
PA (CABL-) CANADIAN BLOOD SERVICES.

XX HU Y, Brown B;
 XX WPI; 2004-766498/75.
 XX
 XX Characterizing a viral infection in a host, for developing treatment for
 XX severe acute respiratory syndrome-coronavirus (SARS-CoV), by determining
 XX homology profile of a viral-based sequence element with an endogenous
 XX host element.
 XX
 XX Example 1; Fig 24; 166pp; English.
 XX
 XX The invention relates to a novel method for characterising a viral
 XX infection in a host. The method comprises identifying at least one viral-
 XX based sequence element in a biological sample obtained from the host,
 XX determining a homology profile of the viral-based sequence element with
 XX at least one endogenous host element and characterising the viral
 XX infection based on the homology profile, where the homology profile is
 XX indicative of a viral behaviour of the viral infection in the host. The
 XX method of the invention demonstrates virucide applications and may be
 XX useful for preparing a medicament for detecting and/or treating a viral
 XX infection or related condition, such as an autoimmune disease e.g. type
 XX II cryoglobulinemia, or lymphoproliferative disorder. The viral
 XX infection may be due to HCV (Hepatitis C virus), HIV or a member of a
 XX Retroviridae, Flaviviridae, Herpesviridae, Papillomaviridae or
 XX Coronaviridae virus family. Treatment of the infection may utilise
 XX vaccine or gene therapy. The target compound of the invention may be
 XX useful for detecting the presence of a virus in a biological sample or
 XX for manufacturing a medicament for treating SARS coronavirus (SARS-CoV)
 XX and/or Human T-lymphotropic virus 1 (HTLV-I) infection. The methods are
 XX further useful for developing treatment regimes to target genotype-
 XX specific viral variants. The current sequence is that of a human
 XX immunoglobulin light chain kappa (IgVlk) protein of the invention.
 XX
 XX Sequence 93 AA;
 XX
 XX Query Match 18.4%; Score 48; DB 8; Length 93;
 XX Best Local Similarity 100.0%; Pred. No. 7.1e-32;
 XX Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 173 WFOQPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
 XX |||||
 XX Db 40 WFOQPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 87
 XX
 XX RESULT 4
 XX AAE35907
 XX ID AAE35907 standard; protein; 100 AA.
 XX AC AAE35907;
 XX DT 17-JUN-2003 (first entry)
 XX DE Human A17 antibody light chain germline protein.
 XX KW Human; cytotoxic T lymphocyte antigen-4; therapy; CTLA-4 antibody;
 XX cancer.
 XX OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX Region 24..39
 XX /note= "Complementarity determining region (CDR) 1"
 XX Region 55..61
 XX /note= "Complementarity determining region (CDR) 2"
 XX Region 94..100
 XX /note= "Complementarity determining region (CDR) 3"
 XX
 XX EP1262193-A1.
 XX
 XX 04-DEC-2002.
 XX
 XX 23-MAY-2002; 2002EP-00253652.

XX 23-MAY-2001; 2001US-0293042P.
 XX (PFIZ) PFIZER PROD INC.
 XX
 XX Hanson DC, Mueller EB;
 XX WPI; 2003-131215/13.
 XX
 XX Use of human anti-cytotoxic T lymphocyte antigen-4 antibody in the
 XX preparation of medicament for the treatment of cancer.
 XX
 XX Disclosure; Fig 7; 76pp; English.
 XX
 XX The invention relates to the use of human anti-cytotoxic T lymphocyte
 XX antigen-4 (anti-CTLA-4) antibody in the preparation of a medicament for
 XX the treatment of cancer such as bone cancer, pancreatic cancer, skin
 XX cancer, cutaneous or intraocular malignant melanoma, uterine and ovarian
 XX cancer, cancer of the anal region, stomach cancer, breast cancer,
 XX testicular cancer, uterine cancer and carcinoma of the fallopian tubes.
 XX The present sequence is human A10/A26 antibody light chain germline
 XX protein. This sequence is used in the invention
 XX
 XX Sequence 100 AA;
 XX
 XX Query Match 18.4%; Score 48; DB 6; Length 100;
 XX Best Local Similarity 100.0%; Pred. No. 7.5e-32;
 XX Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 173 WFOQPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
 XX |||||
 XX Db 40 WFOQPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 87
 XX
 XX RESULT 5
 XX ABO27142
 XX ID ABO27142 standard; protein; 100 AA.
 XX AC ABO27142;
 XX DT 10-SEP-2003 (first entry)
 XX DE Human germline light chain variable region gene segment #22.
 XX KW Human; light chain variable region; VK; humanised antibody;
 XX chimeric antibody; complementarity determining region; CDR;
 XX canonical CDR structure type.
 XX OS Homo sapiens.
 XX US2003039649-A1.
 XX 27-FEB-2003.
 XX
 XX 12-JUL-2002; 2002US-00194975.
 XX 12-JUL-2001; 2001US-0305111P.
 XX (FOOT/) FOOTE J.
 XX Foote J;
 XX WPI; 2003-492151/46.
 XX Making humanized antibody for converting antibody, by making chimeric
 XX antibodies containing complementarity determining region from non-human
 XX antibody and appropriate framework sequences of human antibodies.
 XX Example 1; Fig 2; 31pp; English.
 XX The invention describes a method of making a humanised antibody,
 XX comprising making chimeric antibodies containing a complementarity
 XX determining region (CDR) from a non-human antibody and appropriate

framework sequences (I) of human antibodies. (I) is selected by using canonical CDR structure types of non-human antibody in comparison to germline canonical CDR structure types of human antibodies as the basis for selection, for humanisation. The method is useful for making a humanised antibody or a converted antibody. The method is applicable for converting a subject antibody sequence of any subject species to a less immunogenic form suitable for use in an object species. The method is reliable for identifying suitable human framework sequences to support non-human CDR regions and to provide humanised antibodies that retain high antigen binding with low immunogenicity in humans, without the need for direct comparison of framework sequences, without the need for determining critically important amino acid residues in the framework, and without the need for multiple iteration and construction to obtain humanised antibodies with suitable therapeutic properties. The antibody has high affinity and low immunogenicity without need for comparing framework sequences between non-human and human antibodies. This sequence represents a human light chain variable region gene segment used in the creation of humanised antibodies

Sequence 100 AA;

Query Match 18.4%; Score 48; DB 6; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPGQSPRLIYKVSNRDGVDPFRSGSGTDTLTKISRVEAED 220
|||||
Db 40 WFOQRPGQSPRLIYKVSNRDGVDPFRSGSGTDTLTKISRVEAED 87

RESULT 6

ADFI10189
ID ADFI10189 standard; protein; 100 AA.

AC ADFI10189;

DT 12-FEB-2004 (first entry)

DE Antibody light chain variable region VLK_2-30.

KW Antibody; stability; solubility; antigen binding affinity;
KW variable region; human.

OS Homo sapiens.

PN WO2003074679-A2.

PD 12-SEP-2003.

PF 03-MAR-2003; 2003WO-US006598.

PR 01-MAR-2002; 2002US-0360843P.

PR 29-MAY-2002; 2002US-0384197P.

XX (XENC-) XENCOR.

PI Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;

DR WPI; 2003-722066/68.

PT Computer optimization of physicochemical properties of antibodies
PT comprises analyzing the interactions of amino acids at variable
PT positions.

PS Example 16; Fig 40b; 135pp; English.

The present invention relates to a method for optimizing at least one physico-chemical property of an antibody by a computational screening method. The method comprises: receiving a template antibody structure; selecting at least one variable position belonging to the antibody structure; selecting at least one amino acid to be considered at the variable position(s); analyzing the interaction of each selected amino acid at each variable position with at least part of the remainder of the

CC antibody, including the selected amino acids at other variable positions;
CC and identifying a set of at least one antibody sequence with at least one
CC optimized physico-chemical property. The method is useful for optimizing
CC the physico-chemical properties of an antibody, especially the stability,
CC solubility, or antigen binding affinity. The optimized antibody may be
CC useful for treating a patient. The present sequence is an antibody
CC variable region sequence used to illustrate the invention.

XX Sequence 100 AA;

Query Match 18.4%; Score 48; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPGQSPRLIYKVSNRDGVDPFRSGSGTDTLTKISRVEAED 220
|||||
Db 40 WFOQRPGQSPRLIYKVSNRDGVDPFRSGSGTDTLTKISRVEAED 87

RESULT 7

ADFI10087

ID ADFI10087 standard; protein; 100 AA.

AC ADFI10087;

DT 12-FEB-2004 (first entry)

DE VEGF antibody light chain variable region VLK_2-30.

KW Antibody; stability; solubility; antigen binding affinity;
KW variable region; human; VEGF.

OS Homo sapiens.

PN WO2003074679-A2.

PD 12-SEP-2003.

PF 03-MAR-2003; 2003WO-US006598.

PR 01-MAR-2002; 2002US-0360843P.

PR 29-MAY-2002; 2002US-0384197P.

XX (XENC-) XENCOR.

PI Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;

DR WPI; 2003-722066/68.

PT Computer optimization of physicochemical properties of antibodies
PT comprises analyzing the interactions of amino acids at variable
PT positions.

PS Example 6; Fig 16b; 135pp; English.

The present invention relates to a method for optimizing at least one physico-chemical property of an antibody by a computational screening method. The method comprises: receiving a template antibody structure; selecting at least one variable position belonging to the antibody structure; selecting at least one amino acid to be considered at the variable position(s); analyzing the interaction of each selected amino acid at each variable position with at least part of the remainder of the antibody, including the selected amino acids at other variable positions; and identifying a set of at least one antibody sequence with at least one optimized physico-chemical property. The method is useful for optimizing the physico-chemical properties of an antibody, especially the stability, solubility, or antigen binding affinity. The optimized antibody may be useful for treating a patient. The present sequence is an antibody variable region sequence used to illustrate the invention.

XX Sequence 100 AA;

Query Match 18.4%; Score 48; DB 7; Length 100;

Best Local Similarity 100.0%; Pred. No. 7.5e-32; Mismatches 0; Indels 0; Gaps 0;
Matches 48; Conservative 0;

QY 173 WFOQPGQSPRLIYKVSNRDSGVDPFRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQPGQSPRLIYKVSNRDSGVDPFRFSGSGGTDFTLKISRVEAED 87

RESULT 8

ADFO9982
ID ADF09982 standard; protein; 100 AA.

XX AC ADF09982;

XX DT 12-FEB-2004 (first entry)

XX DE Antibody light chain variable region VLK_2-30.

XX KW Antibody; stability; solubility; antigen binding affinity;
XX variable region; human.

XX OS Homo sapiens.

XX PN WO2003074679-A2.

XX PD 12-SEP-2003.

XX PF 03-MAR-2003; 2003WO-US006598.

XX PR 01-MAR-2002; 2002US-0360843P.

XX PR 29-MAY-2002; 2002US-0384197P.

XX PA (XENC-) XENCOR.

XX PI Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;

XX DR WPI; 2003-722066/68.

XX PT Computer optimization of physicochemical properties of antibodies
XX comprises analyzing the interactions of amino acids at variable
XX positions.

XX PS Disclosure; Fig 2b; 135pp; English.

XX The present invention relates to a method for optimizing at least one
XX physico-chemical property of an antibody by a computational screening
XX method. The method comprises: receiving a template antibody structure;
XX selecting at least one variable position belonging to the antibody
XX structure; selecting at least one amino acid to be considered at the
XX variable position(s); analyzing the interaction of each selected amino
XX acid at each variable position with at least part of the remainder of the
XX antibody, including the selected amino acids at other variable positions;
XX and identifying a set of at least one antibody sequence with at least one
XX optimized physico-chemical property. The method is useful for optimizing
XX the physico-chemical properties of an antibody, especially the stability,
XX solubility, or antigen binding affinity. The optimized antibody may be
XX useful for treating a patient. The present sequence is an antibody
XX variable region sequence used to illustrate the invention.

XX SQ Sequence 100 AA;

Query Match 18.4%; Score 48; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32; Mismatches 0; Indels 0; Gaps 0;
Matches 48; Conservative 0;

QY 173 WFOQPGQSPRLIYKVSNRDSGVDPFRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQPGQSPRLIYKVSNRDSGVDPFRFSGSGGTDFTLKISRVEAED 87

RESULT 9

ADJ80262
ID ADJ80262 standard; protein; 100 AA.

XX AC ADJ80262;

XX DT 06-MAY-2004 (first entry)

XX DE Vkappa gene locus antibody amino acid sequence #22.

XX KW hybrid antibody; antibody; framework region; homology; immunogenicity.

XX OS Homo sapiens.

XX PN WO2003048321-A2.

XX PD 12-JUN-2003.

XX PF 03-DEC-2002; 2002WO-US038450.

XX PR 03-DEC-2001; 2001US-0336591P.

XX PA (ALEX-) ALEXION PHARM INC.

XX PI Rother R, Wu D;

XX DR WPI; 2003-513753/48.

XX PT Producing a hybrid antibody or hybrid antibody fragment by operatively
XX linking the selected framework sequences to one or more complementarity
XX determining regions of the initial antibody.

XX PS Disclosure; SEQ ID NO 22; 77pp; English.

XX The invention relates to a method of producing a hybrid antibody or
XX hybrid antibody fragment by: (i) providing an initial antibody having
XX specificity for a target; (ii) determining the sequence of a variable
XX region of the initial antibody; (iii) selecting a first component of the
XX variable region consisting of FR1, FR2, FR3 and FR4; (iv) comparing the
XX sequence of the first component to sequences contained in a reference
XX database of antibody sequences or antibody fragment sequences from a
XX target species; (v) selecting a sequence from an antibody in the database
XX which demonstrates a high degree of homology to the first component; (vi)
XX selecting a second component of the variable region which is different
XX than the first component, the second component selected from the group
XX consisting of FR1, FR2, FR3 and FR4; (vii) comparing the sequence of the
XX second component to sequences contained in a reference database of
XX antibody sequences or antibody fragment sequences from the target species
XX; (viii) selecting a sequence from the database which demonstrates a high
XX degree of homology to the second component and which is from a different
XX antibody than the selected antibody; and (ix) operatively linking the
XX selected framework sequences to one or more complementarity determining
XX regions (CDRs) of the initial antibody to produce a hybrid antibody or
XX hybrid antibody fragment. The method is useful for producing a hybrid
XX antibody or hybrid antibody fragment (claimed). The antibody and
XX fragments are useful for therapeutic and diagnostic purposes. The method
XX uses entire framework regions from a single antibody variable heavy or
XX variable light chain to receive the CDRs. This produces antibodies that
XX are highly homologous and exhibit reduced immunogenicity while
XX maintaining an optimum binding profile. This sequence represents the
XX amino acid sequence of an antibody from the Vkappa gene locus.

XX SQ Sequence 100 AA;

Query Match 18.4%; Score 48; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32; Mismatches 0; Indels 0; Gaps 0;
Matches 48; Conservative 0;

QY 173 WFOQPGQSPRLIYKVSNRDSGVDPFRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQPGQSPRLIYKVSNRDSGVDPFRFSGSGGTDFTLKISRVEAED 87

RESULT 10

ADO07331
ID ADO07331 standard; protein; 100 AA.

XX ADO07331;
AC 15-JUL-2004 (first entry)
DT Human antibody A17 light chain variable region.
DE Catalytic antibody; human; antibody; gene therapy; protease.
KW Homo sapiens.
OS
PH Location/Qualifiers
FT 1. .23
FT /label= FR1
FT /note= "Framework region 1"
FT 1
FT /note= "Possible Asp component of catalytic triad"
FT 24. .39
FT /label= CDR1
FT /note= "Complementarity determining region 1"
FT 25
FT /note= "Possible Ser component of catalytic triad"
FT 26
FT /note= "Possible Ser component of catalytic triad"
FT 28
FT /note= "Possible Ser component of catalytic triad"
FT 32
FT /note= "Possible Ser component of catalytic triad"
FT 33
FT /note= "Possible Asp component of catalytic triad"
FT 40. .54
FT /label= FR2
FT /note= "Framework region 2"
FT 55. .61
FT /label= CDR2
FT /note= "Complementarity determining region 2"
FT 57
FT /note= "Possible Ser component of catalytic triad"
FT 60
FT /note= "Possible Asp component of catalytic triad"
FT 61
FT /note= "Possible Ser component of catalytic triad"
FT 62. .93
FT /label= FR3
FT /note= "Framework region 3"
FT 94. .100
FT /label= CDR3
FT /note= "Complementarity determining region 3"
FT 98
FT /note= "Possible His component of catalytic triad"
FT
FT WO2004033658-A2.
XX
XX 22-APR-2004.
XX
XX 09-OCT-2003; 2003WO-US032214.
XX
XX 10-OCT-2002; 2002US-0417979P.
XX
XX (INTE-) INTEGRIGEN INC.
XX
XX Smider V, Larrick JW;
XX
XX WPI; 2004-340921/31.
XX
XX Novel recombinant catalytic polypeptide useful for cleaving target
XX proteins or for treating or preventing cancers, comprises a human
XX antibody light chain operably joined to a heterologous antibody heavy
XX chain.
XX
XX Disclosure; Fig 3; 92pp; English.
XX
XX The present sequence is the light chain variable region of human antibody

CC A17. This is one of a repertoire of human kappa light chain sequences
CC ADO07310-ADO07349 screened for putative catalytic triads. Several genes
CC encoding such light chains ADO07282-ADO07309 were cloned for use in
CC recombinant catalytic polypeptides of the invention. These comprise a
CC human antibody light chain operably joined to a heterologous antibody
CC heavy chain. The light chain has a serine protease dyad and endopeptidase
CC activity, and the heavy chain has a predetermined specificity for a
CC target protein. By joining 2 heterologous human antibody chains, one of
CC which supplies the catalytic activity to hydrolyse polypeptides and the
CC other the binding specificity for a target protein, the invention
CC provides for the construction of a repertoire of proteases with
CC customised protein substrate specificities of potentially unlimited
CC number and thus makes possible the effective treatment and/or prevention
CC of any medical condition attributable to the presence or overexpression
CC of an identified protein. The invention also provides nucleic acids
CC encoding the catalytic antibodies (which can be used for gene therapy),
CC host cells, transgenic non-human animals, and methods of cleaving a
CC target protein (in vitro or in vivo) using a recombinant catalytic
CC polypeptide. It also provides a library of recombinant catalytic
CC polypeptides with altered enzymatic activity, and a method of altering
CC the enzymatic activity of the recombinant catalytic polypeptides by
CC mutating at least one complementarity determining region of the heavy
CC chain. A serine protease triad was identified in the present sequence,
CC which can be used as the light chain sequence in catalytic polypeptides
CC of the invention ADO07283.
XX
XX Sequence 100 AA;
SQ
Query Match 18.4%; Score 48; DB 8; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32; Mismatches 0; Gaps 0;
Matches 48; Conservative 0; Indels 0;
QY 173 WFQRPQSGSPRLIYKVSNRDVGVPDRFSGSGGTDTLTKISRVEAED 220
DB 40 WFQRPQSGSPRLIYKVSNRDVGVPDRFSGSGGTDTLTKISRVEAED 87
RESULT 11
ADY75436
ID ADY75436 standard; peptide; 100 AA.
XX
AC ADY75436;
XX
DT 02-JUN-2005 (first entry)
XX
DE Human germline light chain kappa V minigene VKII peptide #3.
XX
KW Antibody engineering; antibody; antibody production; gene library;
KW DNA recombination; gene amplification; primer extension;
KW light chain variable region.
XX
OS Homo sapiens.
XX
XX WO2005023993-A2.
XX
XX 17-MAR-2005.
XX
XX 09-SEP-2004; 2004WO-US029617.
XX
XX 09-SEP-2003; 2003US-0501073P.
XX
XX (INTE-) INTEGRIGEN INC.
XX
XX Sharma V, Leonard L, Smider V;
XX
XX WPI; 2005-223364/23.
XX
XX Producing polynucleotide encoding human germline antibody V-region for
XX generating full-length antibody germline V-region genes, by obtaining V
XX or J minigene and joining V minigene with J minigene, or joining J
XX minigene with V minigene.
XX
XX Disclosure; Fig 12; 52pp; English.

XX The present invention relates to producing germline antibody genes by a
CC completely in vitro approach that mimics the natural process of V(D)J
CC recombination. The antibody genes are completely human and native in
CC their sequence, and libraries of such antibody genes can be constructed
CC which represent an unselected population representing the entire antibody
CC repertoire. The method uses gene amplification to produce a V minigene,
CC and a hybrid primer capable of hybridizing to a V minigene and either a D
CC or V minigene. The hybrid primer facilitates recombination of a V
CC minigene to a D or J minigene to produce a full length V-region gene.
CC Also disclosed is a library comprising member polynucleotides encoding a
CC exogenously rearranged human germline antibody V-regions. In producing a
CC polynucleotide encoding a human germline antibody V-region, a D minigene
CC is further joined to the 3' end of the V minigene and the 5' end of the J
CC minigene. The V minigene or the J minigene in is obtained by chemical
CC synthesis or by amplification from a germline DNA library. Joining the V
CC minigene with at least one J minigene is performed by primer extension
CC using at least two or three oligonucleotide primers. The V minigene is
CC derived from human immunoglobulin kappa locus, human immunoglobulin
CC lambda locus, or human immunoglobulin heavy chain locus. The V-region
CC also comprises a serine protease triad. The human germline antibodies can
CC be used as precursors to more high affinity antibodies, and are useful in
CC the generation of efficiently pairing libraries of heavy and light
CC chains. The present sequence is a human germline light chain kappa region
CC peptide.

XX Sequence 100 AA;

Query Match 18.4%; Score 48; DB 9; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQPGQSPRLIYKVNDRSGVDPDRFSGSGGTDTFLKISRVEAD 220
DB 40 WFOQPGQSPRLIYKVNDRSGVDPDRFSGSGGTDTFLKISRVEAD 87

RESULT 12

AEBl3589
ID AEB13589 standard; protein; 100 AA.

AC AEB13589;

DT 25-AUG-2005 (first entry)

DE Human variable heavy chain region protein, VH_3-43.

XX protein engineering; immunogenicity; germ cell;
KW heavy chain variable region; antibody.

XX Homo sapiens.

OS WO2005056759-A2.

PN 23-JUN-2005.

PD 03-DEC-2004; 2004WO-US040694.

PF 04-DEC-2003; 2003US-0527167P.

PR 21-JUN-2004; 2004US-0581613P.

PR 13-AUG-2004; 2004US-0601665P.

PR 16-OCT-2004; 2004US-0619483P.

XX (XENC-) XENCOR INC.

XX Lazar GA, Desjarlais JR, Hammond PW;

XX WPI; 2005-458579/46.

XX Generating variant protein for host, by comparing parent protein sequence
PT with natural protein sequences from host, analyzing and substituting
PT amino acids of parent sequences with corresponding amino acid string of
PT natural sequence.

XX

PS Disclosure; Fig 1a; 137pp; English.

XX The invention relates to a novel method for generating a variant protein
CC for a host. The method involves comparing a parent protein sequence with
CC natural protein sequences from a host, analyzing the amino acid strings
CC of the parent sequence with corresponding amino acid strings of each of
CC the natural protein sequences, and substituting amino acids of parent
CC protein sequences with corresponding amino acid strings of a natural
CC protein sequence on an amino acid string. The method is useful for
CC generating a variant protein, e.g. a variant antibody for a host as
CC compared to a parent protein. The method enables the generation of
CC variant proteins having an increased host string content and reduced
CC immunogenicity. This sequence represents a variable heavy chain of a
CC human antibody protein used in a human germ line comparison of the
CC invention.

XX Sequence 100 AA;

Query Match 18.4%; Score 48; DB 9; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQPGQSPRLIYKVNDRSGVDPDRFSGSGGTDTFLKISRVEAD 220
DB 40 WFOQPGQSPRLIYKVNDRSGVDPDRFSGSGGTDTFLKISRVEAD 87

RESULT 13

AEBl3642
ID AEB13642 standard; protein; 100 AA.

AC AEB13642;

DT 25-AUG-2005 (first entry)

DE Human variable light kappa chain region protein, Vlk_2-30.

XX protein engineering; immunogenicity; germ cell;
KW light chain variable region; antibody.

XX Homo sapiens.

OS WO2005056759-A2.

PN 23-JUN-2005.

PD 03-DEC-2004; 2004WO-US040694.

PF 04-DEC-2003; 2003US-0527167P.

PR 21-JUN-2004; 2004US-0581613P.

PR 13-AUG-2004; 2004US-0601665P.

PR 16-OCT-2004; 2004US-0619483P.

XX (XENC-) XENCOR INC.

XX Lazar GA, Desjarlais JR, Hammond PW;

XX WPI; 2005-458579/46.

XX Generating variant protein for host, by comparing parent protein sequence
PT with natural protein sequences from host, analyzing and substituting
PT amino acids of parent sequences with corresponding amino acid string of
PT natural sequence.

XX Disclosure; Fig 1a; 137pp; English.

XX The invention relates to a novel method for generating a variant protein
CC for a host. The method involves comparing a parent protein sequence with
CC natural protein sequences from a host, analyzing the amino acid strings
CC of the parent sequence with corresponding amino acid strings of each of
CC the natural protein sequences, and substituting amino acids of parent
CC protein sequences with corresponding amino acid strings of a natural

CC protein sequence on an amino acid string. The method is useful for
 CC generating a variant protein, e.g. a variant antibody for a host as
 CC compared to a parent protein. The method enables the generation of
 CC variant proteins having an increased host string content and reduced
 CC immunogenicity. This sequence represents a variable light chain of a
 CC human antibody protein used in a human germ line comparison of the
 CC invention.

XX
 SQ Sequence 100 AA;

Query Match 18.4%; Score 48; DB 9; Length 100;
 Best Local Similarity 100.0%; Pred. No. 7.5e-32;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQRPGQSPRLIYKVSNRDGVDPFRSGSGGDTFTLKISRVEAED 220
 |||||
 Db 40 WFOQRPGQSPRLIYKVSNRDGVDPFRSGSGGDTFTLKISRVEAED 87
 |||||

RESULT 14
 AED04308
 ID AED04308 standard; protein; 100 AA.
 XX
 AC AED04308;
 XX
 DT 01-DEC-2005 (first entry)
 XX
 DE Human anti-CTLA-4 antibody, kappa chain protein SEQ ID NO:36.
 XX
 KW cancer; cytostatic; neoplasm; antibody; antibody therapy; igg; CTLA-4;
 KW cytotoxic T lymphocyte antigen-4.
 XX
 OS Homo sapiens.
 XX
 KW WO2005092380-A2.
 PN
 XX
 PD 06-OCT-2005.
 XX
 XX 14-MAR-2005; 2005WO-IB000671.
 XX
 XX 26-MAR-2004; 2004US-0556801P.
 XX
 XX (PFIZ) PFIZER PROD INC.
 XX
 XX Gomez-Navarro J, Hanson DC, Mueller EE, Noe DA;
 XX
 XX WPI; 2005-684091/70.
 XX
 XX Treating cancer in mammal who has undergone stem cell transplantation,
 XX comprises administering human anti-cytotoxic T lymphocyte antigen-4
 XX antibody to the mammal.
 XX
 XX Disclosure; SEQ ID NO 36; 135pp; English.

XX The invention relates to a method of treating (M1) cancer in a mammal
 CC comprising administration of a human anti-CTLA-4 (cytotoxic T lymphocyte
 CC antigen-4) antibody. Also included are: treating cancer in a mammal by
 CC administering more than 10 mg/ml of a human anti-CTLA-4 antibody; and
 CC treating cancer in a mammal by administering an effective amount of a
 CC human anti-CTLA-4 antibody to a mammal who has undergone stem cell
 CC transplantation. The cancer is chosen from breast cancer including
 CC metastatic breast cancer, lung cancer including small-cell lung cancer,
 CC bone cancer, pancreatic cancer, skin cancer, cancer of the head or neck,
 CC melanoma including coetaneous or interlobular malignant melanoma, uterine
 CC cancer, ovarian cancer, rectal cancer, cancer of the anal region, stomach
 CC cancer, colon cancer, testicular cancer, uterine cancer, carcinoma of the
 CC fallopian tubes, carcinoma of the endometrium, carcinoma of the cervix,
 CC carcinoma of the vagina, carcinoma of the vulva, Hodgkin's disease, non-
 CC Hodgkin's lymphoma, cancer of the esophagus, cancer of the small
 CC intestine, cancer of the endocrine system, cancer of the thyroid gland,
 CC cancer of the parathyroid gland, cancer of the adrenal gland, sarcoma of
 CC soft tissue, cancer of the urethra, cancer of the penis, prostate cancer,
 CC chronic acute leukemias including acute myeloid leukemia, chronic myeloid

CC leukemia, acute lymphoblastic leukemia, chronic lymphocytic leukemia,
 CC solid tumors of childhood, lymphocytic lymphomas, cutaneous T cell
 CC lymphoma, cancer of the bladder, cancer of the kidney or ureter, renal
 CC cell carcinoma, carcinoma of the renal pelvis, neoplasm of the central
 CC nervous system (CNS), primary CNS lymphoma, tumor angiogenesis, spinal
 CC axis tumor, brain stem glioma, pituitary adenoma, Kaposi's sarcoma,
 CC epidermoid cancer, squamous cell cancer, T-cell lymphoma, environmentally
 CC induced cancers including those induced by asbestos, myeloma,
 CC neuroblastoma, and pediatric sarcomas. The present sequence represents
 CC human anti-CTLA-4 antibody, kappa chain protein SEQ ID NO:36.

XX
 SQ Sequence 100 AA;

Query Match 18.4%; Score 48; DB 9; Length 100;
 Best Local Similarity 100.0%; Pred. No. 7.5e-32;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQRPGQSPRLIYKVSNRDGVDPFRSGSGGDTFTLKISRVEAED 220
 |||||
 Db 40 WFOQRPGQSPRLIYKVSNRDGVDPFRSGSGGDTFTLKISRVEAED 87
 |||||

RESULT 15
 AED06937
 ID AED06937 standard; protein; 100 AA.
 XX
 AC AED06937;
 XX
 DT 01-DEC-2005 (first entry)
 XX
 DE Human antibody V kappa protein sequence, SEQ ID NO:107.
 XX
 KW antibody therapy; antibody engineering; transgenic animal;
 KW drug screening; leukemia; hematological disease; neoplasm; cytostatic;
 KW autoimmune disease; immunosuppressive; immune disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO2005092926-A2.
 PN
 XX
 PD 06-OCT-2005.
 XX
 XX 17-MAR-2005; 2005WO-US009306.
 XX
 XX 19-MAR-2004; 2004US-0554372P.
 PR
 XX 24-MAY-2004; 2004US-0574661P.
 XX
 XX (ABGE-) ABGENIX INC.
 XX
 XX Kellermann S, Green LL, Korver W;
 XX
 XX WPI; 2005-664987/68.
 XX
 XX Selecting an antibody with a reduced risk of inducing a human anti-human
 XX antibody response comprises providing an immunoglobulin gene encoding a
 XX candidate antibody.
 XX
 XX Disclosure; SEQ ID NO 107; 235pp; English.

XX The new invention relates to selecting an antibody with a reduced risk of
 CC inducing an anti-human antibody (HAHA) response in the host. The method
 CC comprises providing an immunoglobulin gene encoding a candidate antibody
 CC and comparing it with an immunoglobulin gene from the recipient. Also
 CC claimed is a method of excluding an antibody from use in the treatment of
 CC a host. Also given are transgenic animals and methods for identifying
 CC antibodies or for screening or agents that inhibit the induction of HAHA
 CC response; kit for detecting an antibody or assessing the risk of a HAHA
 CC response; a method of determining a risk that an antibody will induce
 CC HAHA response in a patient; and a method of increasing the probability
 CC that a HAHA response will be detected in a transgenic mouse. The method
 CC further comprises repeating the steps of providing, comparing, and
 CC selecting for more than one immunoglobulin gene or for every
 CC immunoglobulin V gene of the candidate antibody. The immunoglobulin gene

CC is a V gene, e.g. a VH (heavy) gene or a VL (light) gene. The methods are
CC useful for selecting an antibody with a reduced risk of inducing HAMA
CC response. The antibody, composition, and method are useful for treating
CC diseases, e.g. leukemia or autoimmune disease. The present sequence is a
CC human antibody V kappa protein sequence.
XX

SQ Sequence 100 AA;

Query Match 18.4%; Score 48; DB 9; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGYDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRLIYKVSNRDSGYDPDRFSGSGGTDFTLKISRVEAED 87

Search completed: August 9, 2006, 13:34:15
Job time : 196 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:39:38 ; Search time 49 Seconds
(without alignments)
466.235 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVLVQSGAEVKPGASV.....KVEIKRGLGLVYKDDDK 261

Scoring table:

OLIGO 60.0 , Gapext 60.0

Searched: 650591 seqs, 87530628 residues

Word size : 1

Total number of hits satisfying chosen parameters: 649417

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /EMC Celleria_SIDS3/ptodata/2/iaa/5_COMB.pap.*
 - 2: /EMC Celleria_SIDS3/ptodata/2/iaa/6_COMB.pap.*
 - 3: /EMC Celleria_SIDS3/ptodata/2/iaa/7_COMB.pap.*
 - 4: /EMC Celleria_SIDS3/ptodata/2/iaa/H_COMB.pap.*
 - 5: /EMC Celleria_SIDS3/ptodata/2/iaa/PCTUS_COMB.pap.*
 - 6: /EMC Celleria_SIDS3/ptodata/2/iaa/RE_COMB.pap.*
 - 7: /EMC Celleria_SIDS3/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	48	18.4	100	2 US-09-472-087-113 Sequence 113, App
2	48	18.4	100	2 US-10-194-975-75 Sequence 75, App
3	48	18.4	111	2 US-09-809-739-13 Sequence 13, App
4	48	18.4	111	2 US-09-840-459-11 Sequence 11, App
5	48	18.4	111	2 US-09-840-459-59 Sequence 59, App
6	48	18.4	111	2 US-09-497-625A-11 Sequence 11, App
7	48	18.4	111	2 US-09-497-625A-59 Sequence 59, App
8	48	18.4	112	1 US-07-942-245-28 Sequence 28, App
9	48	18.4	112	1 US-08-477-877B-89 Sequence 89, App
10	48	18.4	112	1 US-08-472-281A-89 Sequence 89, App
11	48	18.4	112	1 US-08-477-989B-89 Sequence 89, App
12	48	18.4	112	2 US-09-840-459-56 Sequence 56, App
13	48	18.4	112	2 US-09-840-459-62 Sequence 62, App
14	48	18.4	112	2 US-09-840-459-66 Sequence 66, App
15	48	18.4	112	2 US-09-497-625A-56 Sequence 56, App
16	48	18.4	112	2 US-09-497-625A-62 Sequence 62, App
17	48	18.4	112	2 US-09-497-625A-66 Sequence 66, App
18	48	18.4	112	2 US-09-254-180C-8 Sequence 8, App
19	48	18.4	112	2 US-09-462-140D-97 Sequence 97, App
20	46	17.6	112	2 US-09-840-459-60 Sequence 60, App
21	46	17.6	112	2 US-09-497-625A-60 Sequence 60, App
22	37	14.2	112	2 US-09-840-459-70 Sequence 70, App
23	37	14.2	112	2 US-09-497-625A-70 Sequence 70, App
24	34	13.0	96	2 US-10-330-613A-54 Sequence 54, App
25	34	13.0	98	2 US-10-194-975-4 Sequence 4, App
26	34	13.0	98	2 US-10-330-613A-53 Sequence 53, App

27 34 13.0 112 2 US-09-840-459-69 Sequence 69, Appl

28 34 13.0 112 2 US-09-497-625A-69 Sequence 69, Appl

29 34 13.0 117 2 US-08-545-809A-105 Sequence 105, App

30 34 13.0 117 2 US-09-515-697-105 Sequence 105, App

31 34 13.0 123 2 US-10-330-613A-21 Sequence 21, Appl

32 34 13.0 288 2 US-09-818-247-22 Sequence 22, Appl

33 34 12.6 112 2 US-09-840-459-67 Sequence 67, Appl

34 34 12.6 112 2 US-09-497-625A-67 Sequence 67, Appl

35 32 12.3 112 2 US-09-840-459-61 Sequence 61, Appl

36 32 12.3 112 2 US-09-497-625A-61 Sequence 61, Appl

37 32 12.3 113 2 US-09-840-459-63 Sequence 63, Appl

38 32 12.3 113 2 US-09-497-625A-63 Sequence 63, Appl

39 31 11.9 138 2 US-09-453-718B-90 Sequence 90, Appl

40 30 11.5 88 2 US-09-254-180C-151 Sequence 151, App

41 30 11.5 112 2 US-09-840-459-58 Sequence 58, Appl

42 30 11.5 112 1 US-08-428-257A-78 Sequence 78, Appl

43 30 11.5 118 2 US-09-199-149-5 Sequence 5, Appl

44 30 11.5 118 2 US-07-987-264-60 Sequence 60, Appl

45 30 11.5 118 2 US-07-987-264-60 Sequence 60, Appl

ALIGNMENTS

RESULT 1

US-09-472-087-113

; Sequence 113, Application US/09472087

; Patent No. 6682736

; GENERAL INFORMATION:

; APPLICANT: HANSON, DOUGLAS C.

; APPLICANT: NEVEU, MARK J.

; APPLICANT: MUELLER, EILLEN E.

; APPLICANT: HANKE, JEFFREY H.

; APPLICANT: GILMAN, STEVEN C.

; APPLICANT: DAVIS, C. GEORFREY

; APPLICANT: CORVALAN, JOSE R.

; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4

; FILE REFERENCE: ABX-PFI

; CURRENT APPLICATION NUMBER: US/09/472,087

; CURRENT FILING DATE: 1999-12-23

; PRIOR APPLICATION NUMBER: 60/113,647

; PRIOR FILING DATE: 1998-12-23

; NUMBER OF SEQ ID NOS: 147

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 113

; LENGTH: 100

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-472-087-113

Query Match 18.4%; Score 48; DB 2; Length 100;

Best Local Similarity 100.0%; Pred. No. 7.8e-30;

Mismatches 0; Indels 0; Gaps 0;

Matches 48; Conservative 0;

QY 173 WFQRRPGSGPRLLIKVSNRDSGVDPFRFGSGSGTDTTLKISRVEAD 220

Db 40 WFQRRPGSGPRLLIKVSNRDSGVDPFRFGSGSGTDTTLKISRVEAD 87

RESULT 2

US-10-194-975-75

; Sequence 75, Application US/10194975

; Patent No. 6881557

; GENERAL INFORMATION:

; APPLICANT: Foote, Jefferson

; TITLE OF INVENTION: Super Humanized Antibodies

; FILE REFERENCE: 501231.01

; CURRENT APPLICATION NUMBER: US/10/194,975

; CURRENT FILING DATE: 2002-10-10

; PRIOR APPLICATION NUMBER: US 60/305,111

; PRIOR FILING DATE: 2001-07-12

; NUMBER OF SEQ ID NOS: 122

; SOFTWARE: PatentIn version 3.1


```
; SEQ ID NO 75
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-194-975-75

Query Match      18.4%; Score 48; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.8e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQRPQSPRLIYKVNRSNDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
    |||||
Db 40 WFOQRPQSPRLIYKVNRSNDSGVDPDRFSGSGGTDTFTLKISRVEAED 87

RESULT 3
US-09-809-739-13
; Sequence 13, Application US/09809739
; Patent No. 6663863
; GENERAL INFORMATION:
; APPLICANT: Horvath, Christopher J.
; APPLICANT: Rao, Patricia E.
; TITLE OF INVENTION: Method of Inhibiting Stenosis and
; TITLE OF INVENTION: Restenosis
; FILE REFERENCE: 1855.1069-003
; CURRENT APPLICATION NUMBER: US/09/809,739
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: US 09/528,267
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-739-13

Query Match      18.4%; Score 48; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 8.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQRPQSPRLIYKVNRSNDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
    |||||
Db 40 WFOQRPQSPRLIYKVNRSNDSGVDPDRFSGSGGTDTFTLKISRVEAED 87

RESULT 4
US-09-840-459-11
; Sequence 11, Application US/09840459
; Patent No. 6696550
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR FILING DATE: 2000-02-03
; PRIOR FILING DATE: 1999-07-22
; PRIOR FILING DATE: 09/121,781
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
```

```
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-459-11

Query Match      18.4%; Score 48; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 8.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQRPQSPRLIYKVNRSNDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
    |||||
Db 40 WFOQRPQSPRLIYKVNRSNDSGVDPDRFSGSGGTDTFTLKISRVEAED 87

RESULT 5
US-09-840-459-59
; Sequence 59, Application US/09840459
; Patent No. 6696550
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-459-59

Query Match      18.4%; Score 48; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 8.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQRPQSPRLIYKVNRSNDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
    |||||
Db 40 WFOQRPQSPRLIYKVNRSNDSGVDPDRFSGSGGTDTFTLKISRVEAED 87

RESULT 6
US-09-497-625A-11
; Sequence 11, Application US/09497625A
; Patent No. 6727349
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-004
; CURRENT APPLICATION NUMBER: US/09/497,625A
; CURRENT FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
```

; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-497-625A-11

Query Match 18.4%; Score 48; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 8.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 173 WFOQRPGSPRLIYKVSNRDSGVDPDRFSGSGTDTFLKISRVEAD 220
|||||
Db 40 WFOQRPGSPRLIYKVSNRDSGVDPDRFSGSGTDTFLKISRVEAD 87

RESULT 7
US-09-497-625A-59
; Sequence 59, Application US/09497625A
; Patent No. 6727349
; GENERAL INFORMATION:
; APPLICANT: Larosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; FILE REFERENCE: 1855.1052-004
; CURRENT APPLICATION NUMBER: US/09/497,625A
; PRIOR FILING DATE: 2000-02-03
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-497-625A-59

Query Match 18.4%; Score 48; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 8.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPGSPRLIYKVSNRDSGVDPDRFSGSGTDTFLKISRVEAD 220
|||||
Db 40 WFOQRPGSPRLIYKVSNRDSGVDPDRFSGSGTDTFLKISRVEAD 87

RESULT 8
US-09-942-245-28
; Sequence 28, Application US/07942245
; Patent No. 5639641
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.
; TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT
; NUMBER OF SEQUENCES: 522
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington

; STATE: D. C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: HP 9000/700 Workstation
; OPERATING SYSTEM: UNIX
; SOFTWARE: In house
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,245
; FILING DATE: 09-SEP-1992
; CLASSIFICATION: 530
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-942-245-28

Query Match 18.4%; Score 48; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 173 WFOQRPGSPRLIYKVSNRDSGVDPDRFSGSGTDTFLKISRVEAD 220
|||||
Db 40 WFOQRPGSPRLIYKVSNRDSGVDPDRFSGSGTDTFLKISRVEAD 87

RESULT 9
US-08-477-877B-89
; Sequence 89, Application US/08477877B
; Patent No. 5730979
; GENERAL INFORMATION:
; APPLICANT: Bazin, Herv
; APPLICANT: Latinne, Dominique
; TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activation
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Carella, Byrnes, Bain, Gilfillan,
; ADDRESSER: Cecchi, Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,877B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,009
; FILING DATE: 29-MAR-1995
; APPLICATION NUMBER: 08/119,032
; FILING DATE: 09-SEP-1993
; APPLICATION NUMBER: 08/027,008
; FILING DATE: 05-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 61750-146
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744

;; INFORMATION FOR SEQ ID NO: 89:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 112 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: polypeptide
;; FEATURE:
;; NAME/KEY: Light chain vairable region of HUM5400
US-08-477-877B-89

Query Match 18.4%; Score 48; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQQRPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
DB 40 WFQQRPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 87

RESULT 10

US-08-472-281A-89
; Sequence 89, Application US/08472281A
; Patent No. 5817311
; GENERAL INFORMATION:
; APPLICANT: Bazin, Herv
; APPLICANT: Latinne, Dominique
; TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activation
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi, Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07068

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.281A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,009
; FILING DATE: 29-MAR-1995
; APPLICATION NUMBER: 08/119,032
; FILING DATE: 09-SEP-1993
; APPLICATION NUMBER: 08/027,008
; FILING DATE: 05-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 61750-142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; FEATURE:
; NAME/KEY: Light chain vairable region of HUM5400
US-08-472-281A-89

Query Match 18.4%; Score 48; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFQQRPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
DB 40 WFQQRPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 87

RESULT 11

US-08-477-989B-89
; Sequence 89, Application US/08477989B
; Patent No. 5951983
; GENERAL INFORMATION:
; APPLICANT: Bazin, Herv
; APPLICANT: Latinne, Dominique
; APPLICANT: Kaplan, Ruth
; APPLICANT: Kieber-Emmons, Thomas
; APPLICANT: Postema, Christina E.
; APPLICANT: White-Scharf, Mary
; TITLE OF INVENTION: LO-CD2a Antibody and Uses
; TITLE OF INVENTION: Thereof for Inhibiting
; TITLE OF INVENTION: T-Cell Activation and
; TITLE OF INVENTION: Proliferation
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi, Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07068

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,989B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,009
; FILING DATE: 29-MAR-1995
; APPLICATION NUMBER: 08/119,032
; FILING DATE: 09-SEP-1993
; APPLICATION NUMBER: 08/027,008
; FILING DATE: 05-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 61750-147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; FEATURE:
; NAME/KEY: Light chain vairable region of
; NAME/KEY: HUM5400
US-08-477-989B-89

Query Match 18.4%; Score 48; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQQRPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
DB 40 WFQQRPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 87

```
RESULT 12
US-09-840-459-56
; Sequence 56, Application US/09840459
; Patent No. 6696550
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 56
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-459-56

Query Match      18.4%; Score 48; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQPGQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFLKISRVEAED 220
Db 40 WFOQPGQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFLKISRVEAED 87

RESULT 13
US-09-840-459-62
; Sequence 62, Application US/09840459
; Patent No. 6696550
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-09-840-459-62

Query Match      18.4%; Score 48; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQPGQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFLKISRVEAED 220
Db 40 WFOQPGQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFLKISRVEAED 87

RESULT 14
US-09-840-459-66
; Sequence 66, Application US/09840459
; Patent No. 6696550
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-459-66

Query Match      18.4%; Score 48; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQPGQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFLKISRVEAED 220
Db 40 WFOQPGQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFLKISRVEAED 87

RESULT 15
US-09-497-625A-56
; Sequence 56, Application US/09497625A
; Patent No. 6727349
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-004
; CURRENT APPLICATION NUMBER: US/09/497,625A
; CURRENT FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 106
```

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 56

; LENGTH: 112

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-497-625A-56

Query Match 18.4%; Score 48; DB 2; Length 112;

Best Local Similarity 100.0%; Pred.No. 8.5e-30;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQPGQSPERLIYKVSNRDVGVPDRFSGSGGTDTFTLKISRVEAED 220

Db 40 WFOQPGQSPERLIYKVSNRDVGVPDRFSGSGGTDTFTLKISRVEAED 87

Search completed: August 9, 2006, 13:41:02

Job time : 50 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:51:33 ; Search time 176 Seconds
(without alignments)
686.927 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVLVQSGAEVKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2097797 seqs, 463214858 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2096646

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA Main.*

- 1: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US08_PUBCOMB.pep.*
- 3: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 4: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 5: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 6: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261	100.0	261	6	US-11-056-825-2
2	217	83.1	259	6	US-11-056-825-7
3	48	18.4	100	4	US-10-194-975-75
4	48	18.4	100	4	US-10-153-382-36
5	48	18.4	100	4	US-10-308-817-22
6	48	18.4	100	4	US-10-453-698-22
7	48	18.4	100	4	US-10-379-392-87
8	48	18.4	100	5	US-10-612-497-113
9	48	18.4	100	5	US-10-776-649-113
10	48	18.4	100	6	US-11-085-368-36
11	48	18.4	100	6	US-11-084-554-107
12	48	18.4	100	6	US-11-054-669-75
13	48	18.4	100	6	US-11-128-900-113
14	48	18.4	100	6	US-11-004-590-82
15	48	18.4	100	6	US-11-136-250-107
16	48	18.4	111	3	US-09-835-087-2
17	48	18.4	111	3	US-09-809-739-13
18	48	18.4	111	3	US-09-840-459-11
19	48	18.4	111	3	US-09-840-459-59
20	48	18.4	111	4	US-10-766-773-11
21	48	18.4	111	4	US-10-766-773-59
22	48	18.4	111	4	US-10-766-610-11
23	48	18.4	111	4	US-10-766-610-59
24	48	18.4	111	4	US-10-733-563-11
25	48	18.4	111	4	US-10-733-563-59
26	48	18.4	111	5	US-10-706-852-15
27	48	18.4	111	5	US-10-662-061-13

```

28 48 18.4 112 3 US-09-840-459-56 Sequence 56, Appl
23 48 18.4 112 3 US-09-840-459-62 Sequence 62, Appl
30 48 18.4 112 3 US-09-840-459-66 Sequence 66, Appl
31 48 18.4 112 4 US-10-766-773-56 Sequence 56, Appl
32 48 18.4 112 4 US-10-766-773-62 Sequence 62, Appl
33 48 18.4 112 4 US-10-766-773-66 Sequence 66, Appl
34 48 18.4 112 4 US-10-766-610-56 Sequence 56, Appl
35 48 18.4 112 4 US-10-766-610-62 Sequence 62, Appl
36 48 18.4 112 4 US-10-766-610-66 Sequence 66, Appl
37 48 18.4 112 4 US-10-733-563-56 Sequence 56, Appl
38 48 18.4 112 4 US-10-733-563-62 Sequence 62, Appl
39 48 18.4 112 4 US-10-733-563-66 Sequence 66, Appl
40 48 18.4 253 3 US-09-880-748-1964 Sequence 1964, Ap
41 48 18.4 253 4 US-10-293-418-1964 Sequence 1964, Ap
42 48 18.4 253 6 US-11-054-515-1964 Sequence 1964, Ap
43 48 18.4 253 6 US-11-266-444-1964 Sequence 1964, Ap
44 46 17.6 112 3 US-09-840-459-60 Sequence 60, Appl
45 46 17.6 112 4 US-10-766-773-60 Sequence 60, Appl

```

ALIGNMENTS

RESULT 1

```

US-11-056-825-2
; Sequence 2, Application US/11056825
; Publication No. US20050255109A1
; GENERAL INFORMATION:
; APPLICANT: Janda, Kim D.
; APPLICANT: Saven, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SGRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-2

```

```

Query Match 100.0%; Score 261; DB 6; Length 261;
Best Local Similarity 100.0%; Pred. No. 7.4e-201;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAQVLVQSGAEVKPGASVKVCASGTFNRYGILTVRQAPGQGLVWGWINGNTHY 60
Db 1 MAQVLVQSGAEVKPGASVKVCASGTFNRYGILTVRQAPGQGLVWGWINGNTHY 60

Qy 61 AQKQGRVTMTDTSTAYMELSLRSDTAVYICARDPRGDDPEYWGQGLTVTVSSGG 120
Db 61 AQKQGRVTMTDTSTAYMELSLRSDTAVYICARDPRGDDPEYWGQGLTVTVSSGG 120

Qy 121 GSGGGGGGGGSEIVLTQSPSLPVLTLQGPASISCRSSQNLVYSDGNTVLSWFQORPGQ 180
Db 121 GSGGGGGGGGSEIVLTQSPSLPVLTLQGPASISCRSSQNLVYSDGNTVLSWFQORPGQ 180

Qy 181 SPRRLIYKVNRRDGVDPDRFGSGSGTDTFTLKISRVEAEDIGVYICMQGTHWPRFTFGQ 240
Db 181 SPRRLIYKVNRRDGVDPDRFGSGSGTDTFTLKISRVEAEDIGVYICMQGTHWPRFTFGQ 240

Qy 241 TKVEIKRGLGLVDYKDDDK 261
Db 241 TKVEIKRGLGLVDYKDDDK 261

```

```
RESULT 2
US-11-056-825-7
; Sequence 7, Application US/11056825
; Publication No. US20050255109A1
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SCRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-7

Query Match      83.1%; Score 217; DB 6; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.3e-165; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

Qy 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSNYGITWVRQAPGQGLFWMGWNNGNTHY 60
Db 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSNYGITWVRQAPGQGLFWMGWNNGNTHY 60

Qy 61 AQKPGQVRVMTTDTSTSTAYNELSLRSDDTAVYICARDPRGDDPEPYWGQGLTVTVSSGG 120
Db 61 AQKPGQVRVMTTDTSTSTAYNELSLRSDDTAVYICARDPRGDDPEPYWGQGLTVTVSSGG 120

Qy 121 GSGGGGGGGGGSEIVLTQSPLSIPVTLTGQPASISCRSSQNLVYSDGNTYLSWFOQRPQ 180
Db 121 GSGGGGGGGGGSEIVLTQSPLSIPVTLTGQPASISCRSSQNLVYSDGNTYLSWFOQRPQ 180

Qy 181 SPRRLIYKVNRRDGVDPDRFSGSGGTDTLTKISRVE 217
Db 181 SPRRLIYKVNRRDGVDPDRFSGSGGTDTLTKISRVE 217

RESULT 3
US-10-194-975-75
; Sequence 75, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-194-975-75

Query Match      18.4%; Score 48; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0;

Qy 173 WFOQRPQSPRLIYKVNRRDGVDPDRFSGSGGTDTLTKISRVEAD 220
Db 173 WFOQRPQSPRLIYKVNRRDGVDPDRFSGSGGTDTLTKISRVEAD 220
```

```
RESULT 4
US-10-153-382-36
; Sequence 36, Application US/10153382
; Publication No. US20030086930A1
; GENERAL INFORMATION:
; APPLICANT: PFIZER PRODUCTS INC.
; TITLE OF INVENTION: USES OF ANTI-CTLA-4 ANTIBODIES
; FILE REFERENCE: PC23019A
; CURRENT APPLICATION NUMBER: US/10/153,382
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: 60/293042
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-382-36

Query Match      18.4%; Score 48; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0;

Qy 173 WFOQRPQSPRLIYKVNRRDGVDPDRFSGSGGTDTLTKISRVEAD 220
Db 40 WFOQRPQSPRLIYKVNRRDGVDPDRFSGSGGTDTLTKISRVEAD 87

RESULT 5
US-10-308-817-22
; Sequence 22, Application US/10308817
; Publication No. US20030219861A1
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; APPLICANT: Wu, Davang
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 1087-37
; CURRENT APPLICATION NUMBER: US/10/308,817
; CURRENT FILING DATE: 2002-12-03
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 100
; TYPE: PRT
; ORGANISM: human
US-10-308-817-22

Query Match      18.4%; Score 48; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0;

Qy 173 WFOQRPQSPRLIYKVNRRDGVDPDRFSGSGGTDTLTKISRVEAD 220
Db 40 WFOQRPQSPRLIYKVNRRDGVDPDRFSGSGGTDTLTKISRVEAD 87

RESULT 6
US-10-453-698-22
; Sequence 22, Application US/10453698
; Publication No. US20040038308A1
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 82 CIP (1087-37 CIP)
; CURRENT APPLICATION NUMBER: US/10/453,698
; CURRENT FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.2
US-10-453-698-22
```

```
; SEQ ID NO 22
; LENGTH: 100
; TYPE: PRT
; ORGANISM: human
; US-10-453-698-22

Query Match      18.4%; Score 48; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGVPDRFSGSGGTDTFLKISRVEAD 220
Db 40 WFOQRPQSPRLIYKVSNRDSGVPDRFSGSGGTDTFLKISRVEAD 87

RESULT 7
US-10-379-392-87
; Sequence 87, Application US/10379392
; Publication No. US20040110226A1
; GENERAL INFORMATION:
; APPLICANT: Lazar, Gregory Alan
; APPLICANT: Desjarlais, John Rudolf
; APPLICANT: Marshall, Shannon Alicia
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION
; FILE REFERENCE: A-71386-3 463077-236
; CURRENT APPLICATION NUMBER: US/10/379,392
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,843
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/384,197
; PRIOR FILING DATE: 2002-05-29
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-379-392-87

Query Match      18.4%; Score 48; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGVPDRFSGSGGTDTFLKISRVEAD 220
Db 40 WFOQRPQSPRLIYKVSNRDSGVPDRFSGSGGTDTFLKISRVEAD 87

RESULT 8
US-10-612-497-113
; Sequence 113, Application US/10612497
; Publication No. US20040228859A1
; GENERAL INFORMATION:
; APPLICANT: Douglas C. Hanson
; APPLICANT: Mark J. Neveu
; APPLICANT: Eileen E. Mueller
; APPLICANT: Jeffrey H. Hanke
; APPLICANT: Steven C. Gilman
; APPLICANT: C. Geoffrey Davis
; APPLICANT: Jose R. Corvalan
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PFI DIV1
; CURRENT APPLICATION NUMBER: US/10/612,497
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: US 09/472087
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 60/113647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 100
; TYPE: PRT
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-612-497-113

Query Match      18.4%; Score 48; DB 5; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGVPDRFSGSGGTDTFLKISRVEAD 220
Db 40 WFOQRPQSPRLIYKVSNRDSGVPDRFSGSGGTDTFLKISRVEAD 87

RESULT 9
US-10-776-649-113
; Sequence 113, Application US/10776649
; Publication No. US20040228861A1
; GENERAL INFORMATION:
; APPLICANT: Douglas C. Hanson
; APPLICANT: Mark J. Neveu
; APPLICANT: Eileen E. Mueller
; APPLICANT: Jeffrey H. Hanke
; APPLICANT: Steven C. Gilman
; APPLICANT: C. Geoffrey Davis
; APPLICANT: Jose R. Corvalan
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PFI DIV2
; CURRENT APPLICATION NUMBER: US/10/776,649
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 10/612497
; PRIOR FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: US 09/472087
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 60/113647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-776-649-113

Query Match      18.4%; Score 48; DB 5; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGVPDRFSGSGGTDTFLKISRVEAD 220
Db 40 WFOQRPQSPRLIYKVSNRDSGVPDRFSGSGGTDTFLKISRVEAD 87

RESULT 10
US-11-085-368-36
; Sequence 36, Application US/11085368
; Publication No. US20050226875A1
; GENERAL INFORMATION:
; APPLICANT: PFIZER PRODUCTS INC.
; APPLICANT: Gomez-Navarro, Jesus
; APPLICANT: Hanson, Douglas C.
; APPLICANT: Eileen, Mueller Elliott
; APPLICANT: Noe, Dennis A.
; TITLE OF INVENTION: USES OF ANTI-CTLA-4 ANTIBODIES
; FILE REFERENCE: PC32177A
; CURRENT APPLICATION NUMBER: US/11/085,368
; CURRENT FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/556,801
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36
; LENGTH: 100
; TYPE: PRT
```


173 WFOORFGOSPRLIYKVSNRD SGVPDRFSGSGGTDFTLKISRVEAED 220

Query Match 18.4%; Score 48; DB 6; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQORPGQSPRRLLYKVSNRDSGVPDRFSGSGGTDTFTLKISRVEAED 220
|||||
Db 40 WFQORPGQSPRRLLYKVSNRDSGVPDRFSGSGGTDTFTLKISRVEAED 87

RESULT 15

US-11-136-250-107
; Sequence 107, Application US/11136250
; Publication No. US20060021074A1
; GENERAL INFORMATION:
; APPLICANT: Kellermann, Sirid-Ai
; APPLICANT: Green, Larry L.
; APPLICANT: Korvet, Wouter
; TITLE OF INVENTION: REDUCING THE RISK OF HUMAN ANTI-HUMAN
; FILE REFERENCE: AGENIX.100A2
; CURRENT APPLICATION NUMBER: US/11/136,250
; CURRENT FILING DATE: 2005-05-23
; PRIOR APPLICATION NUMBER: 11/084,554
; PRIOR FILING DATE: 2005-03-17
; PRIOR APPLICATION NUMBER: PCT/US2005/009306
; PRIOR FILING DATE: 2005-03-17
; PRIOR APPLICATION NUMBER: 60/574,661
; PRIOR FILING DATE: 2004-05-24
; PRIOR APPLICATION NUMBER: 60/554,372
; PRIOR FILING DATE: 2004-03-19
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 107
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-136-250-107

Query Match 18.4%; Score 48; DB 6; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQORPGQSPRRLLYKVSNRDSGVPDRFSGSGGTDTFTLKISRVEAED 220
|||||
Db 40 WFQORPGQSPRRLLYKVSNRDSGVPDRFSGSGGTDTFTLKISRVEAED 87

Search completed: August 9, 2006, 13:54:56
Job time : 177 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:52:08 ; Search time 33 Seconds ;
(without alignments)
532.377 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVLVQSGAEVKKPGASV.....KVEIKRGLGLVYKDDDDK 261

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 239914 seqs, 67312017 residues

Word size: 1

Total number of hits satisfying chosen parameters: 239854

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: Published Applications AA New:

- 1: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 2: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 5: /EMC_Celerra_SIDS3/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 6: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 7: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	18.4	100	7	US-11-239-308-45
2	48	18.4	113	7	US-11-239-308-2
3	35	13.4	247	6	US-10-539-402-18
4	34	13.0	98	7	US-11-221-902-55
5	33	12.6	118	7	US-11-375-221-11
6	31	11.9	32	6	US-10-497-088-7
7	31	11.9	120	7	US-11-337-300-33
8	31	11.9	247	7	US-11-337-300-65
9	31	11.9	291	7	US-11-154-103-10
10	31	11.9	1052	6	US-10-497-088-21
11	31	11.9	1342	6	US-10-497-088-14
12	30	11.5	117	7	US-11-249-296-6
13	30	11.5	453	7	US-11-254-182-44
14	29	11.1	30	6	US-10-570-220-49
15	29	11.1	30	7	US-11-221-902-49
16	29	11.1	87	7	US-11-219-121-19
17	29	11.1	87	7	US-11-061-841-23
18	29	11.1	87	7	US-11-196-917A-19
19	29	11.1	98	7	US-11-221-902-52
20	29	11.1	98	7	US-11-221-902-53
21	29	11.1	98	7	US-11-221-902-54
22	29	11.1	98	7	US-11-221-902-58
23	29	11.1	98	7	US-11-221-902-64
24	29	11.1	108	6	US-10-484-105-14
25	29	11.1	118	7	US-11-249-296-2

26	29	11.1	118	7	US-11-249-296-72	Sequence 72, Appl
27	29	11.1	118	7	US-11-249-296-90	Sequence 90, Appl
28	29	11.1	119	7	US-11-006-808-10	Sequence 10, Appl
29	29	11.1	119	7	US-11-291-140-25	Sequence 25, Appl
30	29	11.1	119	7	US-11-291-140-27	Sequence 27, Appl
31	29	11.1	119	7	US-11-291-140-29	Sequence 29, Appl
32	29	11.1	119	7	US-11-291-140-45	Sequence 45, Appl
33	29	11.1	119	7	US-11-291-140-47	Sequence 47, Appl
34	29	11.1	119	7	US-11-291-140-49	Sequence 49, Appl
35	29	11.1	119	7	US-11-291-140-51	Sequence 51, Appl
36	29	11.1	119	7	US-11-291-140-53	Sequence 53, Appl
37	29	11.1	120	7	US-11-304-986-22	Sequence 22, Appl
38	29	11.1	122	7	US-11-211-917-110	Sequence 110, App
39	29	11.1	126	7	US-11-211-917-42	Sequence 42, Appl
40	29	11.1	129	7	US-11-006-808-45	Sequence 45, Appl
41	29	11.1	139	6	US-10-533-104A-21	Sequence 21, Appl
42	29	11.1	139	6	US-10-533-104A-22	Sequence 22, Appl
43	29	11.1	247	7	US-11-337-300-96	Sequence 96, Appl
44	29	11.1	248	7	US-11-337-300-94	Sequence 94, Appl
45	29	11.1	249	7	US-11-337-300-67	Sequence 67, Appl

ALIGNMENTS

RESULT 1
US-11-239-308-45
; Sequence 45, Application US/11239308
; Publication No. US2006008883A1
; GENERAL INFORMATION:
; APPLICANT: Snider, Vaughn
; APPLICANT: Larrick, James W.
; APPLICANT: Integrigen, Inc.
; TITLE OF INVENTION: Recombinant Catalytic Polypeptides and Their Uses
; FILE REFERENCE: 021216-000310US
; CURRENT APPLICATION NUMBER: US/11/239,308
; PRIOR FILING DATE: 2005-09-28
; PRIOR APPLICATION NUMBER: US/10/683,733
; PRIOR FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: US 60/417,979
; PRIOR FILING DATE: 2002-10-09
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 45
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-239-308-45

Query Match 18.4%; Score 48; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 8.6e-33;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFQRPQSGPRRLIYKVNRSRSGVDFRFGSGSGTDTLKISRVEAD 220
DB 40 WFQRPQSGPRRLIYKVNRSRSGVDFRFGSGSGTDTLKISRVEAD 87

RESULT 2
US-11-239-308-2
; Sequence 2, Application US/11239308
; Publication No. US2006008883A1
; GENERAL INFORMATION:
; APPLICANT: Snider, Vaughn
; APPLICANT: Larrick, James W.
; APPLICANT: Integrigen, Inc.
; TITLE OF INVENTION: Recombinant Catalytic Polypeptides and Their Uses
; FILE REFERENCE: 021216-000310US
; CURRENT APPLICATION NUMBER: US/11/239,308
; CURRENT FILING DATE: 2005-09-28
; PRIOR APPLICATION NUMBER: US/10/683,733
; PRIOR FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: US 60/417,979

; PRIOR FILING DATE: 2002-10-09
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-239-308-2

Query Match 18.4%; Score 48; DB 7; Length 113;
Best Local Similarity 100.0%; Pred. No. 9.4e-33;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQPQSPRLIYKVSNRDVGVPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQPQSPRLIYKVSNRDVGVPDRFSGSGGTDFTLKISRVEAED 87

RESULT 3
US-10-539-402-18
; Sequence 18, Application US/10539402
; Publication No. US20060115477A1
; GENERAL INFORMATION:
; APPLICANT: Xerion Pharmaceuticals AG
; TITLE OF INVENTION: Neuropilin-1 Inhibitor
; FILE REFERENCE: XE12EPC
; CURRENT APPLICATION NUMBER: US/10/539,402
; PRIOR FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US 60/435,893
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: EP 03000615
; PRIOR FILING DATE: 2003-01-15
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 247
; TYPE: PRT
; ORGANISM: human
US-10-539-402-18

Query Match 13.4%; Score 35; DB 6; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 YAKQFQGRVTMTTDTSTSTAYMELRSLRSDDTAVY 94
Db 52 YAKQFQGRVTMTTDTSTSTAYMELRSLRSDDTAVY 86

RESULT 4
US-11-221-902-55
; Sequence 55, Application US/11221902
; Publication No. US20060088522A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: HUMANIZED ANTI-5T4 ANTIBODIES AND ANTI-5T4/CALICHEAMICIN CONJUGATE
; FILE REFERENCE: 040000-0317285
; CURRENT APPLICATION NUMBER: US/11/221,902
; CURRENT FILING DATE: 2005-09-09
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-221-902-55

Query Match 13.0%; Score 34; DB 7; Length 98;
Best Local Similarity 100.0%; Pred. No. 3.6e-21;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 QGRVTMTTDTSTSTAYMELRSLRSDDTAVYCAR 98

Db 65 QGRVTMTTDTSTSTAYMELRSLRSDDTAVYCAR 98

RESULT 5
US-11-375-221-111
; Sequence 111, Application US/11375221
; Publication No. US20060153850A1
; GENERAL INFORMATION:
; APPLICANT: BEDIAN, VAHE
; APPLICANT: DEVALARAJA, MADHAV NARASIMHA
; APPLICANT: FOLTZ, IAN
; APPLICANT: HAAK-FRENSCH, MARY
; APPLICANT: KELLERMANN, SIRID-AIMEE
; APPLICANT: LOW, JOSEPH EDWIN
; APPLICANT: MOBLEY, JAMES LESLIE
; TITLE OF INVENTION: ANTIBODIES TO M-CSF
; FILE REFERENCE: ABX-PP4
; CURRENT APPLICATION NUMBER: US/11/375,221
; CURRENT FILING DATE: 2006-03-13
; PRIOR APPLICATION NUMBER: US/10/938,353
; PRIOR FILING DATE: 2004-09-09
; PRIOR APPLICATION NUMBER: 60/502,163
; PRIOR FILING DATE: 2003-09-10
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 111
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-375-221-111

Query Match 12.6%; Score 33; DB 7; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.8e-20;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 QGRVTMTTDTSTSTAYMELRSLRSDDTAVYCAR 97
Db 65 QGRVTMTTDTSTSTAYMELRSLRSDDTAVYCAR 97

RESULT 6
US-10-497-088-7
; Sequence 7, Application US/10497088
; Publication No. US20060088520A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: Germeaad, Wilfred
; APPLICANT: Logtenberg, Ton
; APPLICANT: Lekkerkerker, Annemarie N
; TITLE OF INVENTION: Antigen presenting cell targeting conjugate, an antigen
; TITLE OF INVENTION: presenting cell contacted with such conjugate, their use for
; TITLE OF INVENTION: vaccination or as medicament, and methods for their production or
; TITLE OF INVENTION: generation
; FILE REFERENCE: 0070 US 00 CON
; CURRENT APPLICATION NUMBER: US/10/497,088
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: PCT/EP01/14255
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: PCT/EP02/13681
; PRIOR FILING DATE: 2002-11-29
; PRIOR APPLICATION NUMBER: EP01204997.9
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: FRI of MatDC16
US-10-497-088-7

```
Query Match      11.9%; Score 31; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 5e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSKASGYTF 31
Db 1 MAQVQLVQSGAEVKKPGASVKVSKASGYTF 31

RESULT 7
US-11-337-300-33
; Sequence 33, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruif, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudmit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT APPLICATION NUMBER: US/11/337.300
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Variable heavy chain of SC03-012
US-11-337-300-33

Query Match      11.9%; Score 31; DB 7; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSKASGYTF 31
Db 1 MAQVQLVQSGAEVKKPGASVKVSKASGYTF 31

RESULT 8
US-11-337-300-65
; Sequence 65, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruif, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudmit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT APPLICATION NUMBER: US/11/337.300
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: SC03-012
US-11-337-300-65

Query Match      11.9%; Score 31; DB 7; Length 247;
Best Local Similarity 100.0%; Pred. No. 2.2e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSKASGYTF 31
Db 2 MAQVQLVQSGAEVKKPGASVKVSKASGYTF 32
```

```
RESULT 9
US-11-154-103-10
; Sequence 10, Application US/11154103
; Publication No. US20060099205A1
; GENERAL INFORMATION:
; APPLICANT: ADAMS, GREGORY P.
; APPLICANT: HORAK, EVA M.
; APPLICANT: WEINER, LOUIS M.
; APPLICANT: JAMES, MARKS D.
; TITLE OF INVENTION: BISPECIFIC SINGLE CHAIN Fv ANTIBODY MOLECULES AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 407T-000420US
; CURRENT APPLICATION NUMBER: US/11/154,103
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US 60/370,276
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US10/406,830
; PRIOR FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic antibody.
US-11-154-103-10

Query Match      11.9%; Score 31; DB 7; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.5e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 111 GTLVTVSSGGSGGGGGGGSEIVLTQSP 141
Db 138 GTLVTVSSGGSGGGGGGGSEIVLTQSP 168

RESULT 10
US-10-497-088-21
; Sequence 21, Application US/10497088
; Publication No. US20060088520A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: Germeraad, Wilfred
; APPLICANT: Logtenberg, Ton
; APPLICANT: Lekkerkerker, Annemarie N
; TITLE OF INVENTION: Antigen presenting cell targeting conjugate, an antigen
; TITLE OF INVENTION: presenting cell contacted with such conjugate, their use for
; TITLE OF INVENTION: vaccination or as medicament, and methods for their production or
; TITLE OF INVENTION: generation
; FILE REFERENCE: 0070 US 00 CON
; CURRENT APPLICATION NUMBER: US/10/497,088
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: PCT/EP01/14255
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: PCT/EP02/13681
; PRIOR FILING DATE: 2002-11-29
; PRIOR APPLICATION NUMBER: EP01204997.9
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 1052
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: MatDC16-Cgamma4-MAGE-1
; NAME/KEY: misc feature
; LOCATION: (546)..(546)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
```


Search completed: August 9, 2006, 13:55:35
Job time : 34 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:40:20 ; Search time 605 Seconds
(without alignments)
658.736 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVLVQSGAEVKKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 8366291 seqs, 1526956180 residues

Word size : 1

Total number of hits satisfying chosen parameters: 8362699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents AA Main.*

1: /EMC_Celerra_SIDS3/ptodata/2/paa/PCTUS_COMB.pep.*
2: /EMC_Celerra_SIDS3/ptodata/2/paa/US066_COMB.pep.*
3: /EMC_Celerra_SIDS3/ptodata/2/paa/US073_COMB.pep.*
4: /EMC_Celerra_SIDS3/ptodata/2/paa/US074_COMB.pep.*
5: /EMC_Celerra_SIDS3/ptodata/2/paa/US075_COMB.pep.*
6: /EMC_Celerra_SIDS3/ptodata/2/paa/US076_COMB.pep.*
7: /EMC_Celerra_SIDS3/ptodata/2/paa/US077_COMB.pep.*
8: /EMC_Celerra_SIDS3/ptodata/2/paa/US078_COMB.pep.*
9: /EMC_Celerra_SIDS3/ptodata/2/paa/US079_COMB.pep.*
10: /EMC_Celerra_SIDS3/ptodata/2/paa/US080_COMB.pep.*
11: /EMC_Celerra_SIDS3/ptodata/2/paa/US081_COMB.pep.*
12: /EMC_Celerra_SIDS3/ptodata/2/paa/US082_COMB.pep.*
13: /EMC_Celerra_SIDS3/ptodata/2/paa/US083_COMB.pep.*
14: /EMC_Celerra_SIDS3/ptodata/2/paa/US084_COMB.pep.*
15: /EMC_Celerra_SIDS3/ptodata/2/paa/US085_COMB.pep.*
16: /EMC_Celerra_SIDS3/ptodata/2/paa/US086_COMB.pep.*
17: /EMC_Celerra_SIDS3/ptodata/2/paa/US087_COMB.pep.*
18: /EMC_Celerra_SIDS3/ptodata/2/paa/US088_COMB.pep.*
19: /EMC_Celerra_SIDS3/ptodata/2/paa/US089_COMB.pep.*
20: /EMC_Celerra_SIDS3/ptodata/2/paa/US090_COMB.pep.*
21: /EMC_Celerra_SIDS3/ptodata/2/paa/US091_COMB.pep.*
22: /EMC_Celerra_SIDS3/ptodata/2/paa/US092_COMB.pep.*
23: /EMC_Celerra_SIDS3/ptodata/2/paa/US093_COMB.pep.*
24: /EMC_Celerra_SIDS3/ptodata/2/paa/US094_COMB.pep.*
25: /EMC_Celerra_SIDS3/ptodata/2/paa/US095_COMB.pep.*
26: /EMC_Celerra_SIDS3/ptodata/2/paa/US096_COMB.pep.*
27: /EMC_Celerra_SIDS3/ptodata/2/paa/US097_COMB.pep.*
28: /EMC_Celerra_SIDS3/ptodata/2/paa/US098_COMB.pep.*
29: /EMC_Celerra_SIDS3/ptodata/2/paa/US099_COMB.pep.*
30: /EMC_Celerra_SIDS3/ptodata/2/paa/US100_COMB.pep.*
31: /EMC_Celerra_SIDS3/ptodata/2/paa/US101_COMB.pep.*
32: /EMC_Celerra_SIDS3/ptodata/2/paa/US102_COMB.pep.*
33: /EMC_Celerra_SIDS3/ptodata/2/paa/US103_COMB.pep.*
34: /EMC_Celerra_SIDS3/ptodata/2/paa/US104_COMB.pep.*
35: /EMC_Celerra_SIDS3/ptodata/2/paa/US105_COMB.pep.*
36: /EMC_Celerra_SIDS3/ptodata/2/paa/US106_COMB.pep.*
37: /EMC_Celerra_SIDS3/ptodata/2/paa/US107_COMB.pep.*
38: /EMC_Celerra_SIDS3/ptodata/2/paa/US108_COMB.pep.*
39: /EMC_Celerra_SIDS3/ptodata/2/paa/US109_COMB.pep.*
40: /EMC_Celerra_SIDS3/ptodata/2/paa/US110_COMB.pep.*
41: /EMC_Celerra_SIDS3/ptodata/2/paa/US111_COMB.pep.*
42: /EMC_Celerra_SIDS3/ptodata/2/paa/US112_COMB.pep.*
43: /EMC_Celerra_SIDS3/ptodata/2/paa/US113_COMB.pep.*

44: /EMC_Celerra_SIDS3/ptodata/2/paa/US114_COMB.pep.*
45: /EMC_Celerra_SIDS3/ptodata/2/paa/US500_COMB.pep.*
46: /EMC_Celerra_SIDS3/ptodata/2/paa/US501_COMB.pep.*
47: /EMC_Celerra_SIDS3/ptodata/2/paa/US502_COMB.pep.*
48: /EMC_Celerra_SIDS3/ptodata/2/paa/US503_COMB.pep.*
49: /EMC_Celerra_SIDS3/ptodata/2/paa/US504_COMB.pep.*
50: /EMC_Celerra_SIDS3/ptodata/2/paa/US505_COMB.pep.*
51: /EMC_Celerra_SIDS3/ptodata/2/paa/US506_COMB.pep.*
52: /EMC_Celerra_SIDS3/ptodata/2/paa/US507_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	261	100.0	261	1	PCT-US05-04612-2	Sequence 2, Appli
2	261	100.0	261	40	US-11-056-825-2	Sequence 2, Appli
3	217	83.1	259	1	PCT-US05-04612-7	Sequence 7, Appli
4	217	83.1	259	40	US-11-056-825-7	Sequence 7, Appli
5	48	18.4	100	1	PCT-US02-22011-75	Sequence 75, Appl
6	48	18.4	100	1	PCT-US02-38450-22	Sequence 22, Appl
7	48	18.4	100	1	PCT-US03-06598-87	Sequence 87, Appl
8	48	18.4	100	31	US-10-153-382-36	Sequence 36, Appl
9	48	18.4	100	33	US-10-308-817-22	Sequence 22, Appl
10	48	18.4	100	33	US-10-379-392-87	Sequence 87, Appl
11	48	18.4	100	34	US-10-453-698-22	Sequence 22, Appl
12	48	18.4	100	36	US-10-612-497-113	Sequence 113, App
13	48	18.4	100	36	US-10-612-497A-113	Sequence 113, App
14	48	18.4	100	36	US-10-683-733-45	Sequence 45, Appl
15	48	18.4	100	37	US-10-776-649-113	Sequence 113, App
16	48	18.4	100	40	US-11-004-590-82	Sequence 82, Appl
17	48	18.4	100	40	US-11-054-669-75	Sequence 75, Appl
18	48	18.4	100	40	US-11-084-554-107	Sequence 107, App
19	48	18.4	100	40	US-11-085-368-36	Sequence 36, Appl
20	48	18.4	100	41	US-11-128-900-113	Sequence 113, App
21	48	18.4	100	41	US-11-136-250-107	Sequence 107, App
22	48	18.4	100	42	US-11-239-308-45	Sequence 45, Appl
23	48	18.4	100	42	US-11-239-308A-45	Sequence 45, Appl
24	48	18.4	104	27	US-09-791-537-107428	Sequence 107428, A
25	48	18.4	108	27	US-09-791-537-23019	Sequence 23019, A
26	48	18.4	110	27	US-09-791-537-24305	Sequence 24305, A
27	48	18.4	110	46	US-60-164-762-847	Sequence 847, App
28	48	18.4	111	1	PCT-US01-03537-11	Sequence 11, Appl
29	48	18.4	111	1	PCT-US01-03537-59	Sequence 59, Appl
30	48	18.4	111	1	PCT-US01-12139-2	Sequence 2, Appli
31	48	18.4	111	1	PCT-US03-39599A-11	Sequence 11, Appl
32	48	18.4	111	1	PCT-US03-39599A-59	Sequence 59, Appl
33	48	18.4	111	28	US-09-835-087-2	Sequence 2, Appli
34	48	18.4	111	36	US-10-662-061-13	Sequence 13, Appl
35	48	18.4	111	37	US-10-766-852-15	Sequence 15, Appl
36	48	18.4	111	37	US-10-733-563-11	Sequence 11, Appl
37	48	18.4	111	37	US-10-733-563-59	Sequence 59, Appl
38	48	18.4	111	37	US-10-766-610-11	Sequence 11, Appl
39	48	18.4	111	37	US-10-766-610-59	Sequence 59, Appl
40	48	18.4	111	37	US-10-766-773-11	Sequence 11, Appl
41	48	18.4	111	37	US-10-766-773-59	Sequence 59, Appl
42	48	18.4	112	1	PCT-US01-03537-56	Sequence 56, Appl
43	48	18.4	112	1	PCT-US01-03537-62	Sequence 62, Appl
44	48	18.4	112	1	PCT-US01-03537-66	Sequence 66, Appl
45	48	18.4	112	1	PCT-US03-39599A-56	Sequence 56, Appl

ALIGNMENTS

RESULT 1

PCT-US05-04612-2
; Sequence 2, Application PC/TUS0504612
; GENERAL INFORMATION:


```
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Sayen, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SGRP-0043
; CURRENT APPLICATION NUMBER: PCT/US05/04612
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: US 11/056,825
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 2
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
PCT-US05-04612-2

Query Match      100.0%; Score 261; DB 1; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.2e-236;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITWVRQAPGQGLEWMGWINNNGNTHY 60
Db      1  MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITWVRQAPGQGLEWMGWINNNGNTHY 60

Qy      61 AQKFGQGRVTMTTDTSTSTAYMELRSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Db      61 AQKFGQGRVTMTTDTSTSTAYMELRSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120

Qy      121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180
Db      121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180

Qy      181 SPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEADIGVYCMQGTTHWPRTFGQ 240
Db      181 SPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEADIGVYCMQGTTHWPRTFGQ 240

Qy      241 TKVEIKRGLGLVDYKDDDK 261
Db      241 TKVEIKRGLGLVDYKDDDK 261

RESULT 2
US-11-056-825-2
; Sequence 2, Application US/11056825
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Sayen, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SGRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 2
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-2
```

```
Query Match      100.0%; Score 261; DB 40; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.2e-236;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITWVRQAPGQGLEWMGWINNNGNTHY 60
Db      1  MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITWVRQAPGQGLEWMGWINNNGNTHY 60

Qy      61 AQKFGQGRVTMTTDTSTSTAYMELRSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Db      61 AQKFGQGRVTMTTDTSTSTAYMELRSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120

Qy      121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180
Db      121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180

Qy      181 SPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEADIGVYCMQGTTHWPRTFGQ 240
Db      181 SPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEADIGVYCMQGTTHWPRTFGQ 240

Qy      241 TKVEIKRGLGLVDYKDDDK 261
Db      241 TKVEIKRGLGLVDYKDDDK 261

RESULT 3
PCT-US05-04612-7
; Sequence 7, Application PC/TUS0504612
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Sayen, Alan
; FILE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SGRP-0043
; CURRENT APPLICATION NUMBER: PCT/US05/04612
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: US 11/056,825
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 7
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
PCT-US05-04612-7

Query Match      83.1%; Score 217; DB 1; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.9e-195;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITWVRQAPGQGLEWMGWINNNGNTHY 60
Db      1  MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITWVRQAPGQGLEWMGWINNNGNTHY 60

Qy      61 AQKFGQGRVTMTTDTSTSTAYMELRSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Db      61 AQKFGQGRVTMTTDTSTSTAYMELRSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120

Qy      121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180
Db      121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180

Qy      181 SPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVE 217
Db      181 SPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVE 217
```

```
RESULT 4
US-11-056-825-7
; Sequence 7, Application US/11056825
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Saven, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SSCP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-7

Query Match      83.1%; Score 217; DB 40; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.9e-195;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSNYGITTWRQAPGGLEWMGWINNNTYH 60
Db 1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSNYGITTWRQAPGGLEWMGWINNNTYH 60

Qy 61 AQKQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Db 61 AQKQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120

Qy 121 GSGGGGGGGGGSEIVLTQSPSLPVLTLGQPASISCRSSQNLVYSDGNTYLSWFOQRPQG 180
Db 121 GSGGGGGGGGGSEIVLTQSPSLPVLTLGQPASISCRSSQNLVYSDGNTYLSWFOQRPQG 180

Qy 181 SPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVE 217
Db 181 SPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVE 217

RESULT 5
PCT-US02-22011-75
; Sequence 75, Application PC/TUS0222011
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.02
; CURRENT APPLICATION NUMBER: PCT/US02/22011
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-22011-75

Query Match      18.4%; Score 48; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFQORPGQSPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVEAED 220
Db 40 WFQORPGQSPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVEAED 87

RESULT 6
PCT-US02-38450-22
; Sequence 22, Application PC/TUS0238450
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 1087-37
; CURRENT APPLICATION NUMBER: PCT/US02/38450
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 100
; TYPE: PRT
; ORGANISM: human
PCT-US02-38450-22

Query Match      18.4%; Score 48; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFQORPGQSPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVEAED 220
Db 40 WFQORPGQSPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVEAED 87

RESULT 7
PCT-US03-06598-87
; Sequence 87, Application PC/TUS0306598
; GENERAL INFORMATION:
; APPLICANT: XENCOR
; APPLICANT: Lazar, Gregory Alan
; APPLICANT: Desjarlais, John Rudolf
; APPLICANT: Marshall, Shannon Alicia
; APPLICANT: Dahiya, Basail I.
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION
; FILE REFERENCE: FP-71386-3 463077-237
; CURRENT APPLICATION NUMBER: PCT/US03/06598
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,843
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/384,197
; PRIOR FILING DATE: 2002-05-29
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-06598-87

Query Match      18.4%; Score 48; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFQORPGQSPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVEAED 220
Db 40 WFQORPGQSPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVEAED 87

RESULT 8
US-10-153-382-36
; Sequence 36, Application US/10153382
; GENERAL INFORMATION:
; APPLICANT: PFIZER PRODUCTS INC.
; TITLE OF INVENTION: USES OF ANTI-CTLA-4 ANTIBODIES
; FILE REFERENCE: PC23019A
; CURRENT APPLICATION NUMBER: US/10/153,382
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: 60/293042
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 39
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-382-36

Query Match      18.4%; Score 48; DB 31; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 9
US-10-308-817-22
; Sequence 22, Application US/10308817
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; APPLICANT: Wu, Dayang
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 1087-37
; CURRENT APPLICATION NUMBER: US/10/308,817
; CURRENT FILING DATE: 2002-12-03
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 100
; TYPE: PRT
; ORGANISM: human
US-10-308-817-22

Query Match      18.4%; Score 48; DB 33; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 10
US-10-379-392-87
; Sequence 87, Application US/10379392
; GENERAL INFORMATION:
; APPLICANT: Lazar, Gregory Alan
; APPLICANT: Desjarlais, John Rudolf
; APPLICANT: Marshall, Shannon Alicia
; APPLICANT: Dahiya, Bassil I.
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION
; FILE REFERENCE: A-71386-3 463077-236
; CURRENT APPLICATION NUMBER: US/10/379,392
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,843
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/384,197
; PRIOR FILING DATE: 2002-05-29
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-379-392-87

Query Match      18.4%; Score 48; DB 33; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 220
```

```
Db 40 WFOQRPQSPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 11
US-10-453-698-22
; Sequence 22, Application US/10453698
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 82 CIP (1087-37 CIP)
; CURRENT APPLICATION NUMBER: US/10/453,698
; CURRENT FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 100
; TYPE: PRT
; ORGANISM: human
US-10-453-698-22

Query Match      18.4%; Score 48; DB 34; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 12
US-10-612-497-113
; Sequence 113, Application US/10612497
; GENERAL INFORMATION:
; APPLICANT: Douglas C. Hanson
; APPLICANT: Mark J. Neveu
; APPLICANT: Eileen E. Mueller
; APPLICANT: Jeffrey H. Hanke
; APPLICANT: Steven C. Gilman
; APPLICANT: C. Geoffrey Davis
; APPLICANT: Jose R. Corvalan
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PFI DIVI
; CURRENT APPLICATION NUMBER: US/10/612,497
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: US 09/472087
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 60/113647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-612-497-113

Query Match      18.4%; Score 48; DB 36; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 13
US-10-612-497A-113
; Sequence 113, Application US/10612497A
; GENERAL INFORMATION:
; APPLICANT: HANSON, DOUGLAS C.
; APPLICANT: NEVEU, MARK J.
; APPLICANT: MUELLER, EILEEN E.
; APPLICANT: HANKE, JEFFREY H.
```

; APPLICANT: GILMAN, STEVEN C.
; APPLICANT: DAVIS, C. GEOFFREY
; APPLICANT: CORVALAN, JOSE R.
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PF1
; CURRENT APPLICATION NUMBER: US/10/612,497A
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: US/09/472,087
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/113,647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-612-497A-113

Query Match 18.4%; Score 48; DB 36; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQPGQSPRRLIYKVSNRDSGVDPDRFSGSGCTDFTLKISRVEAED 220
Db 40 WFOQPGQSPRRLIYKVSNRDSGVDPDRFSGSGCTDFTLKISRVEAED 87

RESULT 14

US-10-683-733-45
; Sequence 45, Application US/10683733
; GENERAL INFORMATION:
; APPLICANT: Smider, Vaughn
; APPLICANT: Larrick, James W.
; APPLICANT: Integrigen, Inc.
; TITLE OF INVENTION: Recombinant Catalytic Polypeptides and Their Uses
; FILE REFERENCE: 021216-000310US
; CURRENT APPLICATION NUMBER: US/10/683,733
; CURRENT FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: US 60/417,979
; PRIOR FILING DATE: 2002-10-09
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-683-733-45

Query Match 18.4%; Score 48; DB 36; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQPGQSPRRLIYKVSNRDSGVDPDRFSGSGCTDFTLKISRVEAED 220
Db 40 WFOQPGQSPRRLIYKVSNRDSGVDPDRFSGSGCTDFTLKISRVEAED 87

RESULT 15

US-10-776-649-113
; Sequence 113, Application US/10776649
; GENERAL INFORMATION:
; APPLICANT: Douglas C. Hanson
; APPLICANT: Mark J. Neveu
; APPLICANT: Eileen E. Mueller
; APPLICANT: Jeffrey H. Hanke
; APPLICANT: Steven C. Gilman
; APPLICANT: C. Geoffrey Davis
; APPLICANT: Jose R. Corvalan
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PF1 Div2
; CURRENT APPLICATION NUMBER: US/10/776,649
; CURRENT FILING DATE: 2004-02-10

; PRIOR APPLICATION NUMBER: US 10/612497
; PRIOR FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: US 09/472087
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 60/113647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-776-649-113

Query Match 18.4%; Score 48; DB 37; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQPGQSPRRLIYKVSNRDSGVDPDRFSGSGCTDFTLKISRVEAED 220
Db 40 WFOQPGQSPRRLIYKVSNRDSGVDPDRFSGSGCTDFTLKISRVEAED 87

Search completed: August 9, 2006, 13:51:13
Job time : 607 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:41:18 ; Search time 35 Seconds
(without alignments)
496.727 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVLVQSGAEVKKPGASV.....KVEIKRGLGLVYKDDDDK 261

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 294602 seqs, 66610880 residues

Word size : 1

Total number of hits satisfying chosen parameters: 293948

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents AA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/paa/US11_NEW_COMB.pep.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	18.4	100	7	US-11-432-734-75
2	48	18.4	100	7	US-11-332-619-87
3	48	18.4	103	7	US-11-346-403A-38
4	48	18.4	111	6	US-10-377-122-12
5	48	18.4	253	8	US-60-797-351-1964
6	48	18.4	253	8	US-60-814-869-1964
7	48	18.4	253	8	US-60-815-559-1964
8	48	18.4	253	8	US-60-834-152-1964
9	39	14.9	246	8	US-60-797-351-1192
10	39	14.9	246	8	US-60-814-869-1192
11	39	14.9	246	8	US-60-815-559-1192
12	39	14.9	246	8	US-60-834-152-1192
13	37	14.2	247	8	US-60-797-351-3242
14	37	14.2	247	8	US-60-814-869-3242
15	37	14.2	247	8	US-60-815-559-3242
16	37	14.2	247	8	US-60-834-152-3242
17	36	13.8	122	7	US-11-411-003-374
18	36	13.8	141	7	US-11-411-003-305
19	36	13.8	255	8	US-60-797-351-1190
20	36	13.8	255	8	US-60-814-869-1190
21	36	13.8	255	8	US-60-815-559-1190
22	36	13.8	255	8	US-60-834-152-1190
23	35	13.4	117	1	PCT-US06-16786-118
24	35	13.4	117	7	US-11-433-924-118
25	35	13.4	118	1	PCT-US06-16786-198

26	35	13.4	118	7	US-11-433-924-198	Sequence 198, App
27	35	13.4	119	1	PCT-US06-16786-406	Sequence 406, App
28	35	13.4	119	7	US-11-411-003-338	Sequence 338, App
29	35	13.4	119	7	US-11-411-003-370	Sequence 370, App
30	35	13.4	119	7	US-11-433-924-406	Sequence 406, App
31	35	13.4	126	1	PCT-US06-05691-4	Sequence 4, Appli
32	35	13.4	138	7	US-11-411-003-321	Sequence 321, App
33	35	13.4	251	8	US-60-797-351-1315	Sequence 1315, Ap
34	35	13.4	251	8	US-60-814-869-1315	Sequence 1315, Ap
35	35	13.4	251	8	US-60-815-559-1315	Sequence 1315, Ap
36	35	13.4	251	8	US-60-834-152-1315	Sequence 1315, Ap
37	35	13.4	253	8	US-60-797-351-1611	Sequence 1611, Ap
38	35	13.4	253	8	US-60-814-869-1611	Sequence 1611, Ap
39	35	13.4	253	8	US-60-815-559-1611	Sequence 1611, Ap
40	35	13.4	254	8	US-60-834-152-1611	Sequence 1611, Ap
41	35	13.4	254	8	US-60-797-351-1779	Sequence 1779, Ap
42	35	13.4	254	8	US-60-797-351-1780	Sequence 1780, Ap
43	35	13.4	254	8	US-60-814-869-1779	Sequence 1779, Ap
44	35	13.4	254	8	US-60-814-869-1780	Sequence 1780, Ap
45	35	13.4	254	8	US-60-815-559-1779	Sequence 1779, Ap

ALIGNMENTS

RESULT 1
US-11-432-734-75
; Sequence 75, Application US/11432734
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: SUPER HUMANIZED ANTIBODIES
; FILE REFERENCE: 30219/US/3
; CURRENT APPLICATION NUMBER: US/11/432,734
; CURRENT FILING DATE: 2006-05-10
; PRIOR APPLICATION NUMBER: US/11/054,669
; PRIOR FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US 10/194,975
; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 75
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-432-734-75

Query Match 18.4%; Score 48; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.4e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQPGSGSPRLIYKVNRSNDRSGVDFRSGSGSGTDTFTLKISRVEAED 220
DB 40 WFOQPGSGSPRLIYKVNRSNDRSGVDFRSGSGSGTDTFTLKISRVEAED 87

RESULT 2
US-11-332-619-87
; Sequence 87, Application US/11332619
; GENERAL INFORMATION:
; APPLICANT: Moore, Gregory L.
; APPLICANT: Marehall, Shannon Alicia
; TITLE OF INVENTION: ANTIBODIES AND FC FUSION PROTEINS WITH ALTERED IMMUNOGENICITY
; FILE REFERENCE: 187350/US/4 463077-427
; CURRENT APPLICATION NUMBER: US/11/332,619
; CURRENT FILING DATE: 2006-01-12
; PRIOR APPLICATION NUMBER: US 60/643,313
; PRIOR FILING DATE: 2005-01-12
; PRIOR APPLICATION NUMBER: US 60/652,958
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: US 60/654,636
; PRIOR FILING DATE: 2005-02-17

; NUMBER OF SEQ ID NOS: 1939
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 87
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-332-619-87

Query Match 18.4%; Score 48; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.4e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFOQRPQGSPPRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQGSPPRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 3
US-11-346-403A-38
; Sequence 38, Application US/11346403A
; GENERAL INFORMATION:
; APPLICANT: Hoet, Rene
; APPLICANT: Schoonbroodt, Sonia
; APPLICANT: Ladner, Robert C.
; TITLE OF INVENTION: LIBRARIES AND METHODS FOR
; TITLE OF INVENTION: ISOLATING ANTIBODIES
; FILE REFERENCE: 10280-116001
; CURRENT APPLICATION NUMBER: US/11/346.403A
; CURRENT FILING DATE: 2006-02-01
; PRIOR APPLICATION NUMBER: US 60/649,065
; PRIOR FILING DATE: 2005-02-01
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-11-346-403A-38

Query Match 18.4%; Score 48; DB 7; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.5e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFOQRPQGSPPRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 41 WFOQRPQGSPPRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 88

RESULT 4
US-10-377-122-12
; Sequence 12, Application US/10377122
; GENERAL INFORMATION:
; APPLICANT: HANSEN, HANS J.
; APPLICANT: LEUNG, SHUI-ON
; APPLICANT: OU, ZHENGXING
; APPLICANT: GOLDENBERG, DAVID M.
; TITLE OF INVENTION: INTERNALIZING ANTI-CD-74 ANTIBODIES AND METHODS OF USE
; FILE REFERENCE: 78258/329656
; CURRENT APPLICATION NUMBER: US/10/377,122
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/360,259
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-377-122-12

Query Match 18.4%; Score 48; DB 6; Length 111;

Best Local Similarity 100.0%; Pred. No. 2.6e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFOQRPQGSPPRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQGSPPRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 5
US-60-797-351-1964
; Sequence 1964, Application US/60797351
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP11
; CURRENT APPLICATION NUMBER: US/60/797,351
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-797-351-1964

Query Match 18.4%; Score 48; DB 8; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.5e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFOQRPQGSPPRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 180 WFOQRPQGSPPRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 227

RESULT 6
US-60-814-869-1964
; Sequence 1964, Application US/60814869
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP12
; CURRENT APPLICATION NUMBER: US/60/814,869
; CURRENT FILING DATE: 2006-06-20
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-814-869-1964

Query Match 18.4%; Score 48; DB 8; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.5e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFOQRPQGSPPRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 180 WFOQRPQGSPPRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 227

RESULT 7
US-60-815-559-1964
; Sequence 1964, Application US/60815559
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP13
; CURRENT APPLICATION NUMBER: US/60/815,559
; CURRENT FILING DATE: 2006-06-22
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-815-559-1964

```

Query Match      18.4%; Score 48; DB 8; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.5e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDGVDPDRFSGSGTDFTLKISRVEAED 220
Db 180 WFOQRPQSPRLIYKVSNRDGVDPDRFSGSGTDFTLKISRVEAED 227

RESULT 8
US-60-834-152-1964
; Sequence 1964, Application US/60834152
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP14
; CURRENT APPLICATION NUMBER: US/60/834,152
; CURRENT FILING DATE: 2006-07-31
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-834-152-1964

Query Match      18.4%; Score 48; DB 8; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.5e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDGVDPDRFSGSGTDFTLKISRVEAED 220
Db 180 WFOQRPQSPRLIYKVSNRDGVDPDRFSGSGTDFTLKISRVEAED 227

RESULT 9
US-60-797-351-1192
; Sequence 1192, Application US/60797351
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP11
; CURRENT APPLICATION NUMBER: US/60/797,351
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1192
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-797-351-1192

Query Match      14.9%; Score 39; DB 8; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.3e-28;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 98
Db 59 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 97

RESULT 10
US-60-814-869-1192
; Sequence 1192, Application US/60814869
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP12
; CURRENT APPLICATION NUMBER: US/60/814,869
; CURRENT FILING DATE: 2006-06-20
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1192
; LENGTH: 246
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
US-60-814-869-1192

Query Match      14.9%; Score 39; DB 8; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.3e-28;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 98
Db 59 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 97

RESULT 11
US-60-815-559-1192
; Sequence 1192, Application US/60815559
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP13
; CURRENT APPLICATION NUMBER: US/60/815,559
; CURRENT FILING DATE: 2006-06-22
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1192
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-815-559-1192

Query Match      14.9%; Score 39; DB 8; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.3e-28;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 98
Db 59 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 97

RESULT 12
US-60-834-152-1192
; Sequence 1192, Application US/60834152
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP14
; CURRENT APPLICATION NUMBER: US/60/834,152
; CURRENT FILING DATE: 2006-07-31
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1192
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-834-152-1192

Query Match      14.9%; Score 39; DB 8; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.3e-28;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 98
Db 59 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 97

RESULT 13
US-60-797-351-3242
; Sequence 3242, Application US/60797351
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP11
; CURRENT APPLICATION NUMBER: US/60/797,351
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 3242

```

```
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-797-351-3242

Query Match      14.2%; Score 37; DB 8; Length 247;
Best Local Similarity 100.0%; Pred. No. 9.7e-27;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      62 QKFGQGVMTTDTSTSTAYMELSLRSLRSDDTAVYYCAR 98
|
Db      62 QKFGQGVMTTDTSTSTAYMELSLRSLRSDDTAVYYCAR 98
|

RESULT 14
US-60-814-869-3242
; Sequence 3242, Application US/60814869
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulators
; FILE REFERENCE: PF523PP12
; CURRENT APPLICATION NUMBER: US/60/814,869
; CURRENT FILING DATE: 2006-06-20
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 3242
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-814-869-3242

Query Match      14.2%; Score 37; DB 8; Length 247;
Best Local Similarity 100.0%; Pred. No. 9.7e-27;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      62 QKFGQGVMTTDTSTSTAYMELSLRSLRSDDTAVYYCAR 98
|
Db      62 QKFGQGVMTTDTSTSTAYMELSLRSLRSDDTAVYYCAR 98
|

RESULT 15
US-60-815-559-3242
; Sequence 3242, Application US/60815559
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulators
; FILE REFERENCE: PF523PP13
; CURRENT APPLICATION NUMBER: US/60/815,559
; CURRENT FILING DATE: 2006-06-22
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 3242
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-815-559-3242

Query Match      14.2%; Score 37; DB 8; Length 247;
Best Local Similarity 100.0%; Pred. No. 9.7e-27;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      62 QKFGQGVMTTDTSTSTAYMELSLRSLRSDDTAVYYCAR 98
|
Db      62 QKFGQGVMTTDTSTSTAYMELSLRSLRSDDTAVYYCAR 98
|

Search completed: August 9, 2006, 13:51:53
Job time : 36 secs
```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:34:32 ; Search time 39 Seconds
(without alignments)
643.912 Million cell updates/sec

Title: US-11-056-825-2
Perfect score: 261
Sequence: 1 MAQVLVQSGAEVKPGASV.....KVEIKRGLGGLVDYKDDDK 261

Scoring table:
OLIGO Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 1

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	48	18.4	114	2 S49572	Ig kappa chain pre
2	48	18.4	114	2 B49002	Ig kappa chain V r
3	48	18.4	120	2 S42268	Ig kappa chain V r
4	48	18.4	133	1 A24452	Ig kappa chain pre
5	48	18.4	133	1 K2HURP	Ig kappa chain pre
6	48	18.4	133	2 S23230	Ig kappa chain pre
7	48	18.4	133	2 S42611	HUNVK protein prec
8	48	18.4	142	2 S22902	Ig kappa chain V r
9	37	14.2	122	2 S40338	Ig kappa chain - h
10	35	13.4	129	2 S36260	Ig heavy chain V r
11	34	13.0	98	2 S26919	Ig heavy chain V r
12	34	13.0	131	2 S21921	Ig heavy chain V r
13	33	12.6	122	2 S36271	Ig heavy chain V r
14	33	12.6	132	2 S40322	Ig kappa chain - h
15	32	12.3	118	2 S40374	Ig kappa chain - h
16	32	12.3	134	2 S19665	Ig heavy chain pre
17	32	12.3	140	2 S22658	Ig heavy chain V r
18	29	11.1	98	2 S26920	Ig heavy chain V r
19	29	11.1	98	2 S26918	Ig heavy chain V r
20	29	11.1	98	2 S26912	Ig heavy chain V r
21	29	11.1	98	2 S26938	Ig heavy chain V r
22	29	11.1	98	2 S26913	Ig heavy chain V r
23	29	11.1	104	2 S69899	Ig heavy chain V r
24	29	11.1	111	2 S21925	Ig heavy chain V r
25	29	11.1	116	2 S31667	Ig heavy chain V r
26	29	11.1	117	1 HVH035	Ig heavy chain pre
27	29	11.1	117	1 HVH035	Ig heavy chain pre
28	29	11.1	117	2 S18552	Ig heavy chain V r
29	29	11.1	117	2 S31680	Ig heavy chain V r

30	29	11.1	117	2 S18553	Ig heavy chain V r
31	29	11.1	117	2 S18551	Ig heavy chain V r
32	29	11.1	118	2 S36265	Ig heavy chain V r
33	29	11.1	129	2 S46393	Ig heavy chain V r
34	29	11.1	132	2 S31596	Ig heavy chain V r
35	29	11.1	135	2 S49530	anti-Sm antibody V
36	29	11.1	142	2 A32483	Ig heavy chain V r
37	29	11.1	148	2 S29257	Ig heavy chain V r
38	28	10.7	120	2 S42267	Ig kappa chain V r
39	27	10.3	54	2 S34093	Ig kappa chain V r
40	27	10.3	71	2 H30538	Ig kappa chain V-J
41	27	10.3	75	2 S40337	Ig kappa chain V-J
42	27	10.3	87	2 S34094	Ig kappa chain V r
43	27	10.3	87	2 S34091	Ig kappa chain V r
44	27	10.3	93	2 PH1039	Ig light chain V r
45	27	10.3	94	2 PL0258	Ig kappa chain V r

ALIGNMENTS

RESULT 1

S49572
Ig kappa chain precursor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Mar-1995 #sequence_revision 14-Jul-1995 #text_change 21-Jan-2000
C:Accession: S49572

R:Giachino, C.; Padova, E.; Lanzavecchia, A.

A:Description: k+l+ dual receptor B cells are present in the human peripheral repertoire.

A:Reference number: S49571

A:Accession: S49572

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-114 <GIA>

A:Cross-references: UNIPARC:UPI0000116709; EMBL:Z46626; NID:G575261; PIDN:CAR86596.1; PII

A:Superfamily: immunoglobulin V region; immunoglobulin homology

F16-95/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.6e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQRPQSGSPRLIYKVSNRDSGVDPFRFSGSGSGTDTLKISRVEAED 220

Db 40 WFQRPQSGSPRLIYKVSNRDSGVDPFRFSGSGSGTDTLKISRVEAED 87

RESULT 2

B49002

Ig kappa chain V region, rheumatoid factor RF antibody - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: B49002

R:Stuber, F.; Lee, S.K.; Bridges Jr., S.L.; Koopman, W.J.; Schroeder, H.W.J.; Gaskin, F.;

Arthritis Rheum. 35, 900-904, 1992

A:Title: A rheumatoid factor from a normal individual encoded by VH2 and V kappa II gene

A:Reference number: A49002; MUID:92352481; PMID:1322670

A:Accession: B49002

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-114 <STU>

A:Cross-references: UNIPROT:Q9UL80; UNIPARC:UPI0000176B43

A:Experimental source: EBV-transformed lymphoblastoid cell line SSH23

A>Note: sequence inconsistent with the nucleotide translation

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F16-95/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.6e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 173 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 40 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 87

RESULT 3
S42268
Ig kappa chain V region (A17) - human
C:Species: Homo sapiens (man)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 31-Dec-2004
C:Accession: S42268
R:Lautner-Rieske, A.; Huber, C.; Weindl, A.; Pargent, W.; Schaeble, K.F.; Thiebe, R.; Z
Eur. J. Immunol. 22, 1023-1029, 1992
A:Title: The human immunoglobulin kappa locus. Characterization of the duplicated A regi
A:Reference number: A49043; MUID:92201291; PMID:1551402
A:Accession: S42268
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <LAU>
A:Cross-references: UNIPROT:Q8TCD0; UNIPARC:UPI0000176EB9; EMBL:X63403
C:Genetics:
A:Introns: 17/1
C:Superfamily: immunoglobulin homology
F:36-115/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.6e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 60 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 107

RESULT 4
A24452
Ig kappa chain precursor V-II region (RPMI 6410) - human
C:Species: Homo sapiens (man)
C>Date: 24-Jan-1988 #sequence_revision 09-Aug-1996 #text_change 16-Jul-1999
C:Accession: A24452
R:Weir, L.; Leder, P.
Nucleic Acids Res. 14, 3957-3970, 1986
A:Title: Structure and expression of a human subgroup II immunoglobulin kappa gene.
A:Reference number: A24452; MUID:86232631; PMID:3086847
A:Accession: A24452
A:Molecule type: DNA
A:Residues: 1-133 <WEI>
A:Cross-references: UNIPARC:UPI0000113B46; GB:M36859; NID:g185932; PIDN:AAA58920.1; PID:
A:Note: this sequence was determined from the differentiated gene
C:Genetics:
A:Gene: GDB:IGKV2
A:Cross-references: GDB:136265
A:Map position: 2p12-2p12
A:Introns: 17/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:21-20/Domain: signal sequence #status predicted <SIG>
F:21-133/Product: Ig kappa chain V-II region (RPMI 6410) #status predicted <MAT>
F:36-115/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.8e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 60 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 107

RESULT 5
K2HURP
Ig kappa chain precursor V-II region (RPMI) - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C:Accession: A01890
R:Klobeck, H.G.; Meindl, A.; Combriato, G.; Solomon, A.; Zachau, H.G.
Nucleic Acids Res. 13, 6499-6513, 1985
A:Title: Human immunoglobulin kappa light chain genes of subgroups II and III.
A:Reference number: A93588; MUID:86041852; PMID:2997711
A:Accession: A01890
A:Molecule type: DNA
A:Residues: 1-133 <KLO>
A:Cross-references: UNIPROT:P06310; UNIPARC:UPI000012E159
A:Note: the sequence was determined from the differentiated gene
C:Genetics:
A:Gene: GDB:IGKV2
A:Cross-references: GDB:136265
A:Map position: 2p12-2p12
A:Introns: 17/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:21-20/Domain: signal sequence #status predicted <SIG>
F:21-133/Product: Ig kappa chain V-II region (RPMI) #status predicted <MAT>
F:36-115/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.8e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 60 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 107

RESULT 6
S23230
Ig kappa chain precursor V-J region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S23230
R:Kennedy, M.A.
J. Exp. Med. 173, 1033-1036, 1991
A:Title: Novel chromosome translocation caused by fusion of immunoglobulin heavy and lig
A:Reference number: S23230; MUID:91178438; PMID:1840606
A:Accession: S23230
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-133 <KEN>
A:Cross-references: UNIPARC:UPI0000115EA9; EMBL:X55400; NID:g33999; PIDN:CAA39072.1; PID:
C:Genetics:
A:Introns: 17/1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:36-115/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.8e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 60 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 107

```

RESULT 7

S42611
HUNK protein precursor - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S42611
R:Spatz, L.A.; Williams, M.; Brender, B.; Desai, R.; Latov, N.
J. Neuroimmunol. 36, 29-39, 1992
A:Title: DNA sequence analysis and comparison of the variable heavy and light chain regions
A:Reference number: S42610; MUID:92138794; PMID:1370957
A:Accession: S42611
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-133 <SPA>
A:Cross-references: UNIPARC:UPI000011378B; EMBL:X54137; NID:9433889; PIDN:CAA38072.1; PID:1370957
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:35-115/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.9e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFOQRPQSPRLIYKVNRSNGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 60 WFOQRPQSPRLIYKVNRSNGVDPDRFSGSGGTDTFTLKISRVEAED 107

RESULT 8

S22902
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S22902
R:Chastagner, P.; Theze, J.; Zouali, M.
Gene 101, 305-306, 1991
A:Title: Cloning of a gene encoding a lupus-associated human autoantibody V(K) region using a cDNA library
A:Reference number: S22902; MUID:91276289; PMID:1905262
A:Accession: S22902
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-142 <CHA>
A:Cross-references: UNIPARC:UPI0000176CAB; EMBL:X56510
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:47-126/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 1.9e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFOQRPQSPRLIYKVNRSNGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 71 WFOQRPQSPRLIYKVNRSNGVDPDRFSGSGGTDTFTLKISRVEAED 118

RESULT 9

S40338
Ig kappa chain - human
C:Species: Homo sapiens (man)
C:Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C:Accession: S40338
R:Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A:Title: Expressed human immunoglobulin chi genes and their hypermutation.
A:Reference number: S40312; MUID:94080891; PMID:8258341
A:Accession: S40338
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-122 <KLE>
A:Cross-references: UNIPARC:UPI0000116158; EMBL:X72448; NID:9441364; PIDN:CAA51116.1; PID:8258341
C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin
F:22-101/Domain: immunoglobulin homology <IMM>

Query Match 14.2%; Score 37; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 2e-26;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 135 IVLTQSPISLPVTLGPASISCRSSQNLVYSDGNTYL 171
Db 8 IVLTQSPISLPVTLGPASISCRSSQNLVYSDGNTYL 44

RESULT 10

S36260
Ig heavy chain V region (clone alpha-CEA4-8A) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999
C:Accession: S36260
R:Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.; WEMO J. 12, 725-734, 1993
A:Title: Human anti-self antibodies with high specificity from phage display libraries.
A:Reference number: S36256; MUID:93178448; PMID:7679990
A:Accession: S36260
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-129 <GRI>
A:Cross-references: UNIPARC:UPI0000118DEB; EMBL:Z18851; NID:933124; PIDN:CAA79303.1; PID:7679990
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 13.4%; Score 35; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.4e-24;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 65 QQRVTMTTDTSTAYMELSLRSLRSDDTAVYYCARD 99
Db 65 QQRVTMTTDTSTAYMELSLRSLRSDDTAVYYCARD 99

RESULT 11

S26919
Ig heavy chain V region (DP-14) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26919
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V(H) sequences
A:Reference number: S26885; MUID:93021117; PMID:1404388
A:Accession: S26919
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <ROM>
A:Cross-references: UNIPARC:UPI0000031F31; EMBL:Z12316; NID:932855; PIDN:CAA78186.1; PID:1404388
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 13.0%; Score 34; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 65 QQRVTMTTDTSTAYMELSLRSLRSDDTAVYYCARD 98
Db 65 QQRVTMTTDTSTAYMELSLRSLRSDDTAVYYCARD 98

RESULT 12

S21924
Ig heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999

C;Accession: S21924; S21923
R;Friedman, D.F.

submitted to the EMBL Data Library, July 1991

A;Reference number: S21923

A;Accession: S21924

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-131 <FRI>

A;Cross-references: UNIPARC:UPI0000115FA2; EMBL:X60505; NID:g33565; PIDN:CAA43025.1; PID

C;Genetics:

A;Introns: 16/1

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;34-117/Domain: immunoglobulin homology <IMM>

Query Match 13.0%; Score 34; DB 2; Length 131;

Best Local Similarity 100.0%; Pred. No. 1.2e-23;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 QGRVTMTTDTSTAYMELRLSRSDDTAVYYCAR 98

|||||

Db 84 QGRVTMTTDTSTAYMELRLSRSDDTAVYYCAR 117

RESULT 13

S36271

Ig heavy chain V region (clone alpha-THY-29) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999

C;Accession: S36271

R;Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.

EMBO J. 12, 725-734, 1993

A;Title: Human anti-self antibodies with high specificity from phage display libraries.

A;Reference number: S36256; MUID:93178448; PMID:767990

A;Accession: S36271

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-122 <GRI>

A;Cross-references: UNIPARC:UPI0000118DE3; EMBL:Z18832; NID:g33115; PIDN:CAA79284.1; PID

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-98/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 100.0%; Pred. No. 9.1e-23;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 QGRVTMTTDTSTAYMELRLSRSDDTAVYYCA 97

|||||

Db 65 QGRVTMTTDTSTAYMELRLSRSDDTAVYYCA 97

RESULT 14

S40322

Ig kappa chain - human

C;Species: Homo sapiens (man)

C;Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000

C;Accession: S40322

R;Klein, R.; Jaenichen, R.; Zachau, H.G.

Eur. J. Immunol. 23, 3248-3271, 1993

A;Title: Expressed human immunoglobulin chi genes and their hypermutation.

A;Reference number: S40312; MUID:94080891; PMID:8258341

A;Accession: S40322

A;Status: preliminary; translation not shown

A;Molecule type: mRNA

A;Residues: 1-132 <KLE>

A;Cross-references: UNIPARC:UPI0000116148; EMBL:X72432; NID:g441332; PIDN:CAA51100.1; PI

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;35-114/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 100.0%; Pred. No. 9.8e-23;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 188 KVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220

|||||

Db 74 KVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 106

RESULT 15

S40374

Ig kappa chain - human

C;Species: Homo sapiens (man)

C;Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 31-Dec-2004

C;Accession: S40374

R;Klein, R.; Jaenichen, R.; Zachau, H.G.

Eur. J. Immunol. 23, 3248-3271, 1993

A;Title: Expressed human immunoglobulin chi genes and their hypermutation.

A;Reference number: S40312; MUID:94080891; PMID:8258341

A;Accession: S40374

A;Status: preliminary; translation not shown

A;Molecule type: mRNA

A;Residues: 1-118 <KLE>

A;Cross-references: UNIPROT:O8TCD0; UNIPARC:UPI0000176C9D; EMBL:X72484

C;Superfamily: immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;20-99/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 100.0%; Pred. No. 7.3e-22;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 VSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220

|||||

Db 60 VSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 91

Search completed: August 9, 2006, 13:40:06

Job time : 40 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:31:07 ; Search time 301 Seconds
(without alignments)
802.090 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVLVQSGAEVKKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2849598 seqs, 925015592 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	18.4	86	2	Q7Z3Y5 HUMAN
2	48	18.4	114	2	Q9UL80 homo sapien
3	48	18.4	133	1	KV2F HUMAN
4	41	15.7	239	2	Q8TCD0 HUMAN
5	33	12.6	500	2	Q6N091 HUMAN
6	32	12.3	159	2	Q96Q80 HUMAN
7	29	11.1	117	1	HV1B HUMAN
8	29	11.1	117	1	HV1G HUMAN
9	27	10.3	112	2	Q53VP8 mouse
10	27	10.3	113	1	KV2D HUMAN
11	27	10.3	113	1	KV2G HUMAN
12	27	10.3	115	2	Q5F210 MOUSE
13	27	10.3	117	1	KV2E HUMAN
14	27	10.3	219	2	Q65ZC0 MOUSE
15	27	10.3	239	2	Q58E08 HUMAN
16	27	10.3	240	2	Q6PIH6 HUMAN
17	27	10.3	248	2	Q65ZQ7 HUMAN
18	26	10.0	244	2	Q65ZC8 HUMAN
19	24	9.2	124	2	Q9UL92 HUMAN
20	24	9.2	150	2	Q9Y298 HUMAN
21	23	8.8	115	1	KV2A HUMAN
22	23	8.8	119	2	Q9UL94 HUMAN
23	23	8.8	125	2	Q9UL95 HUMAN
24	23	8.8	240	2	Q65ZC9 HUMAN
25	23	8.8	519	2	Q5EBM2 HUMAN
26	21	8.0	170	2	Q925S2 MOUSE
27	21	8.0	239	2	Q8NEK0 HUMAN
28	20	7.7	241	2	Q92IA6 HUMAN
29	19	7.3	243	2	Q7TQM2 HUMAN
30	19	7.3	255	2	Q6KB05 HUMAN
31	19	7.3	498	2	Q6N041 HUMAN

32	19	7.3	500	2	Q9BRV0 HUMAN	Q9brv0 homo sapien
33	19	7.3	518	2	Q6N030 HUMAN	Q6n030 homo sapien
34	18	6.9	108	1	KV1 CANFA	P01618 canis famil
35	18	6.9	109	1	KV4D HUMAN	P83593 homo sapien
36	18	6.9	112	1	KV2D MOUSE	P01629 mus musculu
37	18	6.9	114	1	KV4A HUMAN	P01625 homo sapien
38	18	6.9	121	1	KV40 HUMAN	P06312 homo sapien
39	18	6.9	133	1	KV4E HUMAN	P06313 homo sapien
40	18	6.9	134	1	KV4C HUMAN	P06314 homo sapien
41	17	6.5	108	1	KV5P MOUSE	P01649 mus musculu
42	17	6.5	117	1	HV06 MOUSE	P01750 mus musculu
43	17	6.5	497	2	Q8WY24 HUMAN	Q8wy24 homo sapien
44	16	6.1	218	2	Q925S1 MOUSE	Q925s1 mus musculu
45	16	6.1	262	2	Q65Z11 MOUSE	Q65z11 mus musculu

ALIGNMENTS

```
RESULT 1
Q7Z3Y5 HUMAN
ID Q7Z3Y5 HUMAN PRELIMINARY; PRT; 86 AA.
AC Q7Z3Y5;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DE Rearranged VKA17 V gene segment (Fragment).
GN Name=VKAL17;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Hodgkin lymphoma;
RA Tinguely M., Rosenquist R., Sundstroem C., Amini R.M., Kuppers R.,
RA Hansmann M.L., Brauningner A.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AJ564426; CAD92033.1; -; Genomic_DNA.
CC HSSP; P01625; 1EEQ.
CC SMR; Q7Z3Y5; 1-80.
CC Ensembl; ENSG00000173758; Homo sapiens.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003596; IG_v.
CC InterPro; IPR013106; V-set.
CC SMART; SM00406; IGV; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 86
SQ SEQUENCE 86 AA; 9764 MW; D198FC04FE0C78FD CRC64;
Query Match 18.4%; Score 48; DB 2; Length 86;
Best Local Similarity 100.0%; Pred.No. 3.6e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFQRPQSGPRRLIYKVNSRDSGVDFRFGSGSGTGDTFTLKISRVEAD 220
DB 19 WFQRPQSGPRRLIYKVNSRDSGVDFRFGSGSGTGDTFTLKISRVEAD 66
RESULT 2
Q9UL80 HUMAN
ID Q9UL80 HUMAN PRELIMINARY; PRT; 114 AA.
AC Q9UL80;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DE 07-FEB-2006, entry version 21.
DE Myosin-reactive immunoglobulin light chain variable region (Fragment).
```

```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92352481; PubMed=1322670;
RA Stuber F., Lee S.K., Bridges S.L. Jr., Koopman W.J., Schroeder H.W. Jr.,
RA Gaskin F., Fu S.M.;
RA "A rheumatoid factor from a normal individual encoded by VH2 and V
RT kappa II gene segments.";
RL Arthritis Rheum. 35:900-904(1992).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93170387; PubMed=8436174;
RA Wagner S.D., Luzzatto L.;
RA "V kappa gene segments rearranged in chronic lymphocytic leukemia are
RT distributed over a large portion of the V kappa locus and do not show
RT somatic mutation.";
RL Eur. J. Immunol. 23:391-397(1993).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92289816; PubMed=1601042;
RA Huber C., Klobbeck H.G., Zachau H.G.;
RA "Ongoing V kappa-J kappa recombination after formation of a productive
RT V kappa-J kappa coding joint.";
RL Eur. J. Immunol. 22:1561-1565(1992).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AF035034; AAD56270.1; -; mRNA.
DR PIR; B49002; B49002.
DR PIR; S23638; S23638.
DR PIR; S34094; S34094.
DR PIR; S34095; S34095.
DR HSP; P01625; 11VE.
DR SMR; Q9UL80; 1-114.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG v.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS0835; IG LIKE; 1.
DR NON TER 1 114 114
DR NON TER 114 114
SQ SEQUENCE 114 AA; 12775 MW; 070E31E210D1CB01 CRC64;

Query Match 18.4%; Score 48; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 4.7e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 40 WFOQPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 87

RESULT 3
KV2F_HUMAN
ID KV2F_HUMAN STANDARD; PRT; 133 AA.
AC P06310;
DT 01-JAN-1988, integrated into UniProtKB/Swiss-Prot.
```

```
DT 01-JAN-1988, sequence version 1.
DE 07-MAR-2006, entry version 41.
DI Ig kappa chain V-II region RPMI 6410 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=85041852; PubMed=2997711;
RA Klobbeck H.G., Meindl A., Combriato G., Solomon A., Zachau H.G.;
RT "Human immunoglobulin kappa light chain genes of subgroups II and
RT III.";
RL Nucleic Acids Res. 13:6499-6513(1985).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; Z00020; CAA77315.1; -; Genomic_DNA.
DR PIR; A01890; K2HURP.
DR HSP; Q99M37; 1I91.
DR SMR; P06310; 2I-133.
DR Ensembl; ENSG00000173758; Homo sapiens.
DR LinkHub; P06310; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0003823; P:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG v.
DR Pfam; PF07686; V-set; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS0835; IG LIKE; 1.
DR Immunoglobulin domain; Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 133
FT REGION 21 43
FT REGION 44 59
FT REGION 60 74
FT REGION 75 81
FT REGION 82 113
FT REGION 114 122
FT REGION 123 132
FT DISULFID 43 133
FT NON TER 133 133
SQ SEQUENCE 133 AA; 14707 MW; 513CCAF3673009EE CRC64;

Query Match 18.4%; Score 48; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 5.4e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 60 WFOQPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 107

RESULT 4
Q8TCD0_HUMAN
ID Q8TCD0_HUMAN PRELIMINARY; PRT; 239 AA.
AC Q8TCD0;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
```

RA NP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RA NP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RA NP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92285150; PubMed=1598223;
RA Hirabayashi Y., Munakata Y., Sasaki T., Sano H.;
RT "Variable regions of a human anti-DNA antibody O-81 possessing lupus
RT nephritis-associated idiotype.";
RL Nucleic Acids Res. 20:2601-0(1992).
RN [4]
RA NP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92201291; PubMed=1551402;
RA Lautner-Rieske A., Huber C., Meindl A., Pargent W., Schable K.F.,
RA Thiele R., Zocher I., Zachau H.G.;
RT "The human immunoglobulin kappa locus. Characterization of the
RT duplicated A regions.";
RL Eur. J. Immunol. 22:1023-1029(1992).
RN [5]
RA NP NUCLEOTIDE SEQUENCE
RX MEDLINE=94080891; PubMed=8258341;
RA Klein R., Jaenichen R., Zachau H.G.;
RT "Expressed human immunoglobulin kappa genes and their hypermutation.";
RL Eur. J. Immunol. 23:3248-3262(1993).
RN [6]
RA NP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93170387; PubMed=8436174;
RA Wagner S.D., Luzzatto L.;
RT "V kappa gene segments rearranged in chronic lymphocytic leukemia are
RT distributed over a large portion of the V kappa locus and do not show
RT somatic mutation.";
RL Eur. J. Immunol. 23:391-397(1993).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; BC022362; AAH22362.1; -; mRNA.
DR PIR; S22658; S22658.
DR PIR; S34095; S34095.
DR PIR; S40324; S40324.
DR PIR; S40374; S40374.
DR PIR; S42267; S42267.
DR PIR; S42268; S42268.
DR HSSP; P01834; I172.
DR SNR; Q8TCD0; 21-237.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003597; Ig.cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR InterPro; IPR013106; V-set.
DR Pfam; PF07654; C1-set; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00407; IGcl; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS0835; IG LIKE; 2.
DR PROSITE; PS00290; IG MHC; UNKNOWN 1.
KW Hypothetical protein.
SQ SEQUENCE 239 AA; 26235 MW; FAGEDC3A3B03871D CRC64;
Query Match 15.7%; Score 41; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 4,1e-30;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFOQPGQSPRLIYKVNRRDGVDPDFSGSGSGTDTFTLKI 213
DB 60 WFOQPGQSPRLIYKVNRRDGVDPDFSGSGSGTDTFTLKI 100
RESULT 5
Q6N091 HUMAN PRELIMINARY; PRT; 500 AA.
ID Q6N091_HUMAN PRELIMINARY; PRT; 500 AA.
AC Q6N091;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Hypothetical protein DKFP686C02220 (Fragment).
GN Name=DKFP686C02220;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RA NP NUCLEOTIDE SEQUENCE.
RC TISSUE=Human rectum tumor;
RG The German Human CDNA Consortium;
RA Wambutt R., Haubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; BX640625; CAB45779.1; -; mRNA.
DR HSSP; P01751; IAGW.
DR SMR; Q6N091; 270-478.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig.cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR InterPro; IPR013106; V-set.
DR Pfam; PF07654; C1-set; 2.
DR SMART; SM00409; IG; 1.
DR SMART; SM00407; IGcl; 2.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS0835; IG LIKE; 4.
DR PROSITE; PS00290; IG MHC; UNKNOWN 2.
KW Hypothetical protein; Immunoglobulin domain; Repeat.
FT NON_TER 1
SQ SEQUENCE 500 AA; 54160 MW; 3C423A17D65A41B4 CRC64;
Query Match 12.6%; Score 33; DB 2; Length 500;
Best Local Similarity 100.0%; Pred. No. 3,2e-22;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 65 QGRVTMTTDTSTAYMELRLSLRSDDTAVYYCA 97
DB 102 QGRVTMTTDTSTAYMELRLSLRSDDTAVYYCA 134

DR Pfam; PF07686; V-set; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin domain; Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117
FT DOMAIN 20 >117
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 13009 MW; BB61CB63F8CE97BD CRC64;
Query Match 11.1%; Score 29; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 5.3e-19;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTF 31
DB 20 QVQLVQSGAEVKKPGASVKVSCKASGYTF 48

RESULT 9
Q53VP8 MOUSE
ID Q53VP8 MOUSE PRELIMINARY; PRT; 112 AA.
AC Q53VP8;
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 24-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Kappa chain (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE
RX MEDLINE=86136012; PubMed=3937730;
RA Ollier P., Rocca-Serra J., Somme G., Theze J., Fougereau M.;
RT "The idiotypic network and the internal image: possible regulation of
a germ-line network by paucigene encoded Ab2 (anti-idiotypic)
antibodies in the GAT system";
RL EMBO J. 4:3681-3688(1985).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 108-109.
RA Fougereau M.;
RL Submitted (NOV-1986) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; X03386; CAA27113.1; -; mRNA.
DR DR
DR SMART; SM00406; IGV; 1.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_v.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin domain.
FT NON_TER 1 112
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12270 MW; C844B7881A89C18A CRC64;
Query Match 10.3%; Score 27; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 4.1e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 194 SGVDFRFGSGSGTDFTLKISRVEAED 220
DB 61 SGVDFRFGSGSGTDFTLKISRVEAED 87

RESULT 10
KV2D_HUMAN

ID KV2D_HUMAN STANDARD; PRT; 113 AA.
AC P01617;
DT 21-JUL-1986, integrated into UniProtKB/Swiss-Prot.
DT 21-JUL-1986, sequence version 1.
DT 07-MAR-2006, entry version 45.
DE Ig kappa chain V-II region TEW.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP PROTEIN SEQUENCE (BENCE-JONES PROTEIN TEW).
RX MEDLINE=74148480; PubMed=4596149;
RA Putnam F.W., Whitley E.J. Jr., Paul C., Davidson J.N.;
RT "Amino acid sequence of a kappa Bence Jones protein from a case of
primary amyloidosis";
RL Biochemistry 12:3763-3780(1973).
RN [2]
RP PROTEIN SEQUENCE OF 1-27 (AMYLOID PROTEIN TEW).
RX MEDLINE=73166638; PubMed=4700495;
RA Terry W.D., Page D.L., Kimura S., Isobe T., Osseman E.F.,
RA Glenner G.G.;
RT "Structural identity of Bence Jones and amyloid fibril proteins in a
patient with plasma cell dyscrasia and amyloidosis";
RL J. Clin. Invest. 52:1278-1281(1973).
CC -!- MISCELLANEOUS: The major amyloid protein appears to be identical
CC with the Bence Jones protein isolated from the same patient.
CC -!- MISCELLANEOUS: This protein was isolated from the urine of a
CC patient with plasma cell dyscrasia and amyloidosis.
CC -!- MISCELLANEOUS: The C region of this chain has the INV (1,2)
CC marker.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC PIR; A50370; K2HUTW.
DR HSSP; Q99M37; 1191.
DR SMR; P01617; 1-111.
DR LinkHub; P01617; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0003823; P:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_v.
DR InterPro; IPR013106; V-set.
DR Pfam; PF07686; V-set; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Amyloid; Bence-Jones protein; Direct protein sequencing;
KW Immunoglobulin domain; Immunoglobulin V region.
FT CHAIN 1 >113
FT CHAIN 1 113
FT REGION 1 23
FT REGION 24 39
FT REGION 40 54
FT REGION 55 61
FT REGION 62 93
FT REGION 94 102
FT REGION 103 112
FT DISULFID 23 93
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12316 MW; 0C3C38F81F1843CA CRC64;
Query Match 10.3%; Score 27; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 4.1e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY .194 SGVDFRFGSGSGTDFTLKISRVEAED 220
DB 61 SGVDFRFGSGSGTDFTLKISRVEAED 87

```

RESULT 11
KV2G MOUSE STANDARD; PRT; 113 AA.
AC P01631;
DT 21-JUL-1986, integrated into UniProtKB/Swiss-Prot.
DT 21-JUL-1986, sequence version 1.
DE 07-MAR-2006, entry version 39.
DE Ig kappa chain V-II region 26-10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22716456; PubMed=6404298;
RA Novotny J., Margolles M.N.;
RT "Amino acid sequence of the light chain variable region from a mouse
RL Biochemistry 22:1153-1158(1983).
CC -!- MISCELLANEOUS: This chain was isolated from an IgG2a hybridoma
CC protein that binds digoxin.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR PIR; A01914; KVM526.
DR HSP; Q99M37; I191.
DR Ensembl; ENSMUSG00000055315; Mus musculus.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG V.
DR InterPro; IPR013106; V-set.
DR Pfam; PF07686; V-set; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG LIKE; 1.
KW Immunoglobulin V region; Monoclonal antibody.
KW Immunoglobulin V region; Hybridoma; Immunoglobulin domain;
FT CHAIN 1 >113
FT Ig kappa chain V-II region 26-10.
FT /FTID=PRO_0000059776.
FT REGION 1 23
FT REGION 24 39
FT REGION 40 54
FT REGION 55 61
FT REGION 62 93
FT REGION 94 102
FT REGION 103 112
FT REGION 103 112
FT DISULFID 23 93
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12273 MW; F9F39CE949A84C2A CRC64;

Query Match 10.3%; Score 27; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 4.1e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 SGVPRFSGSGGTDFTLKISRVEAED 220
Db 61 SGVPRFSGSGGTDFTLKISRVEAED 87

RESULT 12
Q5F210 MOUSE
ID Q5F210 MOUSE PRELIMINARY; PRT; 115 AA.
AC Q5F210;
DT 15-MAR-2005, integrated into UniProtKB/TrEMBL.
DT 15-MAR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Kappa light chain variable region (Fragment).
DE Name=IgG1 anti-TS1 VL;
CN Name=IgG1 anti-TS1 VL;

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22716456; PubMed=12833571; DOI=10.1002/jmr.617;
RA Erlandsson A., Holm P., Ullen A., Stigbrand T., Sundstrom B.E.;
RT "Studies of the interactions between the anticytokerin 8 monoclonal
RT antibody TS1, its antigen and its anti-idiotypic antibody alphaTS1.";
RL J. Mol. Recognit. 16:157-163(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Erlandsson A.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ884575; CA156337.1; -; mRNA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG V.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG LIKE; 1.
KW Immunoglobulin domain.
FT NON_TER 1 115
FT NON_TER 115 115
SQ SEQUENCE 115 AA; 12560 MW; E4D3BF3D63E88007 CRC64;

Query Match 10.3%; Score 27; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 4.2e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 SGVPRFSGSGGTDFTLKISRVEAED 220
Db 61 SGVPRFSGSGGTDFTLKISRVEAED 87

RESULT 13
KV2E HUMAN
ID KV2E HUMAN STANDARD; PRT; 117 AA.
AC P06309;
DT 01-JAN-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-JAN-1988, sequence version 1.
DT 07-MAR-2006, entry version 43.
DE Ig kappa chain V-II region GM607 precursor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
OC Homo.
NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=84191506; PubMed=6325927;
RA Klobeck H.G., Solomon A., Zachau H.G.;
RT "Contribution of human V kappa II germ-line genes to light-chain
RT diversity.";
RL Nature 309:73-76(1984).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; Z00009; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR PIR; A01889; K2HUGM.
DR HSP; Q99M37; I191.
DR SMR; P06309; 5-115.
DR LinkHub; P06309; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.

```

DR	PROSITE; PS50835; IG-LIKE; 2.
DR	PROSITE; PS00290; IG-MHC; 1.
KW	Immunoglobulin domain; Repeat.
FT	NON_TER 1
FT	NON_TER 219 219
SQ	SEQUENCE 219 AA; 23945 MW; 7E1B82A14EAP8445 CRC64;
 Query Match 10.3%; Score 27; DB 2; Length 219; Best Local Similarity 100.0%; Pred. No. 7.6e-17; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps	
Qy	194 SGVDPFSGSGSDFTLKISRVEAD 220
Dd	61 SGVDPFSGSGSDFTLKISRVEAD 87
 RESULT 15	
Q58EU8_MOUSE	
ID	Q58EU8_MOUSE PRELIMINARY; PRT; 239 AA.
AC	Q58EU8;
CD	26-APR-2005, integrated into UniProtKB/TrEMBL.
DT	26-APR-2005, sequence version 1.
DT	07-FEB-2006, entry version 12.
DE	Igk-C protein.
DN	Name=Igk-C;
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC	Muridea; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=CZECH II; TISSUE=Mammary tumor metastatized to lung.
RC	MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR enhancer.;
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Straussberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,
RA	Klausner R.D., Collins F.S., Wagner I., Shenem C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Bustow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA	Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Faney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA	Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[2]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=CZECH II; TISSUE=Mammary tumor metastatized to lung.
RC	MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR enhancer.;
RG	NIH MGC Project;
RL	Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC	-----
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC	Distributed under the Creative Commons Attribution-NoDerivs License
CC	-----
CC	EMBL; BC091750; AAH91750.1; -; mRNA.
DR	SMR; Q58EU8; 21-239.
DR	MGI; MGI:96495; Igk-C.
DR	InterPro; IPR003599; Ig.
DR	InterPro; IPR007110; Ig-like.
DR	InterPro; IPR003597; Ig cl.
DR	InterPro; IPR003006; Ig MHC.
DR	InterPro; IPR003596; Ig V.

DR InterPro; IPR011106; V-set.
 DR Pfam; PF07654; Cl-set; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00407; IGc1; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG LIKE; 2.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN 1.
 SQ SEQUENCE 239 AA; 26302 MW; 98FC4BA8BB404215 CRC64;

Query Match 10.3%; Score 27; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 8.2e-17;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 SGVPDRFSGSGGTDFTLKISRVEAED 220
 |||||
 Db 81 SGVPDRFSGSGGTDFTLKISRVEAED 107

Search completed: August 9, 2006, 13:39:20
 Job time : 302 secs

XX New scFv Bc-12 or Bc-15 antibody comprises ligand mimetic, which
PT specifically binds to activated alphavbeta3 integrin receptor, useful for
PT treating, e.g. neoplastic disease, solid tumor, leukemia, or benign or
PT malignant breast cancer.
XX
PS Claim 2; SEQ ID NO 2; 109pp; English.
XX
CC The invention describes an antibody comprising a ligand mimetic, which
CC specifically binds to an activated alphavbeta 3 integrin receptor
CC differentially produced on a cell in a metastatic state compared to a
CC similar, non-metastatic cell, and does not bind to a non-activated
CC alphavbeta 3 integrin receptor. Also described are: a pharmaceutical
CC composition comprising the antibody; treating a disease state in a mammal
CC ; a cell line comprising a tumor cell variant with a metastatic homing
CC propensity to a target tissue; producing an antibody phage population
CC having affinity for a tumor cell target; detecting tumor cells in a
CC mammal by treatment with a cancer therapeutic; inducing or enhancing an
CC immune response to an antigen in a mammal; detecting an activated cell
CC surface receptor on a metastatic tumor cell surface in a mammalian tissue
CC sample; interfering with cells liable to undergo metastasis associated
CC with a disease state; identifying cells liable to undergo metastasis
CC associated with a disease state; an isolated Bc-12 or Bc-15
CC polynucleotide comprising a nucleotide sequence that has at least 90%
CC percent identity to a fully defined 802 or 752 bp sequence (SEQ ID NO. 1
CC or 3) given in the specification; an isolated polypeptide comprising a
CC nucleotide sequence that has at least 90% sequence identity to SEQ ID NO.
CC 1 or 3, or shares a biological function with Bc-12 or Bc-15; a vector
CC comprising the polynucleotide; an expression vector comprising the
CC polynucleotide in which the nucleotide sequence of the polynucleotide is
CC operatively linked with a regulatory sequence that controls expression of
CC the polynucleotide in a host cell; a host cell comprising the
CC polynucleotide or progeny of the cell; and determining anti-metastatic
CC activity of a test compound in a mammal. The antibody is useful for
CC treating neoplastic disease, solid tumor, hematological malignancy,
CC leukemia, colorectal cancer, benign or malignant breast cancer, uterine
CC cancer, uterine leiomyomas, ovarian cancer, endometrial cancer,
CC polycystic ovary syndrome, endometrial polyps, prostatic cancer, prostatic
CC hyper trophy, pituitary cancer, adenomyosis, adenocarcinomas, meningioma,
CC melanoma, bone cancer, multiple myeloma, CNS cancer, glioma, or
CC astroblastoma. This is the amino acid sequence of single chain variable
CC fragment antibody Bc-12.
XX
SQ Sequence 260 AA;

Query Match 98.8%; Score 1381; DB 9; Length 260;
Best Local Similarity 99.6%; Pred. No. 3.2e-91;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MAQVQLVQSGAEVKKPKASVKVSKASGYTFSNYGITWVRQAPGQGLEWMGWINNGNTHY 60
Db 2 MAQVQLVQSGAEVKKPKASVKVSKASGYTFSNYGITWVRQAPGQGLEWMGWINNGNTHY 61
Qy 61 AQKPFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDDPYWGQGLTVTVSSGG 120
Db 62 YQKPFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDDPYWGQGLTVTVSSGG 121
Qy 121 GSGGGGGGGGSEIVLTQSLPLSVTLGQPASISCRSSQNLVYSDGNTYLSWFOORPQ 180
Db 122 GSGGGGGGGGSEIVLTQSLPLSVTLGQPASISCRSSQNLVYSDGNTYLSWFOORPQ 181
Qy 181 SPRRLIYKVNDRSGVDPDRFSGSGSGTFTLKISRVEAEDIGVYCMQGTWPPRPTFGQ 240
Db 182 SPRRLIYKVNDRSGVDPDRFSGSGSGTFTLKISRVEAEDIGVYCMQGTWPPRPTFGQ 241
Qy 241 TKVEIKRGLGLVDYKDD 259
Db 242 TKVEIKRGLGLVDYKDD 260

RESULT 2
ID ABP45953 standard; protein; 253 AA.

XX ABP45953;
AC 19-AUG-2002 (first entry)
XX
XX Human Blys binding scFv SEQ ID 1964.
XX
KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.
XX
OS Homo sapiens.
XX
XX WO200202641-A1.
XX
XX 10-JAN-2002.
XX
XX 15-JUN-2001; 2001WO-US019110.
XX
XX 16-JUN-2000; 2000US-0212210P.
XX
XX 17-OCT-2000; 2000US-0240816P.
XX
XX 16-MAR-2001; 2001US-0276248P.
XX
XX 21-MAR-2001; 2001US-0277379P.
XX
XX 25-MAY-2001; 2001US-0293499P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX (CAME-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
XX WPI; 2002-114799/15.
XX
XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
XX diagnosis and treatment of cancers and immune disorders.
XX
XX Claim 1; Page 2744-2745; 3148pp; English.
XX
XX This invention describes novel antibodies that immunospecifically bind to
XX B Lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
XX tumour necrosis factor (TNF) super family and induces B cell
XX proliferation and differentiation. The antibodies of the invention have
XX cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
XX antirheumatic, and antiAIDS activity and can be used in vaccines to
XX inhibit the expression and activity of Blys. The antibodies bind to Blys
XX and so may be used to detect and quantitate the presence of Blys in
XX biological samples and may be used in this way to diagnose disease
XX associated with aberrant expression of Blys. They may also be
XX administered to treat diseases associated with aberrant Blys expression
XX and activity such as cancer, immune, and autoimmune disorders and
XX diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
XX immunodeficiency (e.g. common variable immunodeficiency (CVID) and
XX acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
XX the antibodies and fragments of the antibodies described in the method of
XX the invention
XX
SQ Sequence 253 AA;

Query Match 81.2%; Score 1135; DB 5; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.4e-73;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;
Qy 3 QVQLVQSGAEVKKPKASVKVSKASGYTFSNYGITWVRQAPGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLVQSGAEVKKPKASVKVSKASGYTFSNYGITWVRQAPGQGLEWMGWIN--NGNTHY 60
Qy 61 AQKPFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDDPYWGQGLTV 115
Db 61 AQKPFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDDPYWGQGLTV 120
Qy 116 VSSGGGGGGGGGGGG--EIVLTQSLPLSVTLGQPASISCRSSQNLVYSDGNTYLSW 173

[illegible][illegible]

Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

DR WPI; 2005-808635/82.
 XX New antibodies that immunospecifically binds to B Lymphocyte Stimulator
 XX protein, useful for diagnosing, treating, or preventing autoimmune
 PT disease, e.g. systemic lupus erythematosus or rheumatoid arthritis, or B
 PT cell cancer.
 XX
 PS Claim 1; SEQ ID NO 1964; 240pp; English.
 XX
 CC The invention relates to a novel antibody that immunospecifically binds
 CC to a B Lymphocyte Stimulator protein. The protein comprises an amino acid
 CC sequence that is 85% identical to the VH domain of any one of the single-
 CC chain variable fragments (scFvs) of SEQ ID NOS. 1-2128, and/or an amino
 CC acid sequence that is at least 85% identical to the VL domain of any one
 CC of the scFvs of SEQ ID NOS. 1-2128. The invention further comprises: an
 CC isolated nucleic acid molecule encoding the antibody; an isolated cell
 CC line that expresses the antibody; a method for detecting the expression
 CC of a B Lymphocyte Stimulator protein; a method for diagnosing an
 CC autoimmune disease or a B cell cancer; and a method for treating,
 CC preventing, or ameliorating an autoimmune disease or a B cell cancer. The
 CC antibody is useful for detecting expression of B Lymphocyte Stimulator
 CC protein, and in diagnosing, treating, preventing, or ameliorating an
 CC autoimmune disease or a B cell cancer. The autoimmune disease is systemic
 CC lupus erythematosus or rheumatoid arthritis. It can also be used for
 CC diagnosing, treating, and preventing immune disorders (e.g. multiple
 CC sclerosis, myasthenia gravis, or Hashimoto's disease), inflammatory
 CC disorders (e.g. asthma or allergic disorders), infectious diseases (e.g.
 CC AIDS), and proliferative disorders (e.g. leukemia, carcinoma, or
 CC lymphoma). This sequence represents a single-chain variable fragment
 CC polypeptide that immunospecifically binds to a B Lymphocyte Stimulator
 CC protein of the invention. Note: This sequence is not shown in the
 CC specification. It has been electronically downloaded from the USPTO
 CC website.
 XX
 SQ Sequence 253 AA;

Query Match 81.2%; Score 1135; DB 9; Length 253;
 Best Local Similarity 86.2%; Pred. No. 1.4e-73;
 Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;
 Qy 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYGTSVVRQAPGQGLEWMGWIN--NGNTHY 60
 Db 1 QVQLQSGAEVKKPGASVKVSCKASGYTFTSYGTSVVRQAPGQGLEWMGWISAYNGNTNY 60
 Qy 61 AOKQGRVTMTTDTSTAYMELSLRSDDTAVYVCARDPRGDD- ----PWGQGLTAVT 115
 Db 61 AOKLQGRVTMTTDTSTAYMELSLRSDDTAVYVCARNPYYSSESFFDYWGQGTMTVT 120
 Qy 116 VSSGGSGSGSGSGSGSGS--EIVLTQSPSLPVLTPGQSPASISCRSSQNLVYSDGNTYLSW 173
 Db 121 VSSGGSGSGSGSGSGSGSALDVMTQSPSLPVLTPGQSPASISCRSSQSLVYSDGNTYLNW 180
 Qy 174 FQQRPGQSPRLIYKVSNRDGVDPFRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
 Db 181 FQQRPGQSPRLIYKVSNRDGVDPFRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239
 Qy 234 PRTFGQTKVIEKR 247
 Db 240 PFTFGQGTMEIKR 253

RESULT 5
 AAY21884
 ID AAY21884 standard; protein; 242 AA.
 XX
 AC AAY21884;
 XX
 DT 24-SEP-1999 (first entry)
 XX
 DE Amino acid sequence of antibody BIOC7.
 XX
 KW prostate-specific antigen; PSA; anti-chymotrypsin; ACT; antibody;
 KW diagnostic assay; PSA-ACT; prostate cancer; tumour; immunotherapy; BIOC7.

XX Homo sapiens.
 XX Key Location/Qualifiers
 FH Misc-difference 102 /label= unknown
 FT /note= "encoded by AST"
 FT Misc-difference 114 /label= unknown
 FT /note= "encoded by TCN"
 FT Misc-difference 155 /label= unknown
 FT /note= "encoded by TGT"
 PN EP934953-A2.
 PD 11-AUG-1999.
 XX 01-DEC-1998; 98EP-00122546.
 XX 03-DEC-1997; 97US-0067428P.
 XX (BOEF) BOEHRINGER MANNHEIM CORP.
 XX Winter GP, Mahoney W, Sawyer JR;
 DR WPI; 1999-432068/37.
 DR N-PSDB; AAX86942.
 FT New anti-complex antibody useful for diagnosing prostate cancer.
 XX Claim 29; Page 28-30; 42pp; English.
 CC The invention relates to an antibody that binds a complex between
 CC prostate-specific antigen (PSA) and anti-chymotrypsin (ACT), and has an
 CC affinity for the complex which is at least 10 fold higher than the
 CC affinity for either PSA or ACT alone. The antibody is used in diagnostic
 CC assays to detect PSA-ACT in serum samples from patients. The levels of
 CC PSA-ACT complex increase in patients suffering from prostate cancer
 CC compared to the levels in patients with benign or no growths in the
 CC prostate. Therefore detection of PSA in complex with ACT is useful for
 CC the early detection of prostate tumours, by distinguishing between benign
 CC and malignant conditions of the prostate as well as for the management of
 CC patients with prostate cancer, such as the disclosure of metastasis and
 CC the monitoring of the PSA levels after treatment. The antibodies may also
 CC be used in immunotherapy, affinity chromatography and isolation or
 CC purification of PSA-ACT. Unlike prior art antibodies which bind to PSA-
 CC ACT complexes but may also bind PSA or ACT alone, the present antibody is
 CC specific for PSA-ACT alone. Diagnostic assays using the antibodies are
 CC more accurate in diagnosing prostate cancer as they only detect the
 CC intact complex of PSA-ACT. Sequences AAY21880-884 represent specific
 CC examples of antibodies directed against PSA-ACT. The present sequence
 CC represents the amino acid sequence of antibody BIOC7
 XX Sequence 242 AA;

Query Match 72.4%; Score 1011.5; DB 2; Length 242;
 Best Local Similarity 78.7%; Pred. No. 1e-64;
 Matches 196; Conservative 20; Mismatches 24; Indels 9; Gaps 3;
 Qy 1 MAQVLVQSGAEVKKPGASVKVSCKASGYTFTSYGTSVVRQAPGQGLEWMGWIN--NGNT 58
 Db 1 MAQVLVQSGAEVKKPGASVKVSCKASGYTFTSYGTSVVRQAPGQGLEWMGWISAYNGNT 60
 Qy 59 HYAQKFGQGRVTMTTDTSTAYMELSLRSDDTAVYVCARDPRGDDPEYWGQGLTAVTSS 118
 Db 61 KYSQKFGQGRVTITRDTSTASTAYMELSLRSDDTAVYVCARGXR- ----FWGQGLTAVTSS 115
 Qy 119 GGGSGSGSGSGSGSGSEIVLTQSPSLPVLTPGQSPASISCRSSQNLVYSDGNTYLSWFQORP 178
 Db 116 GGGSGSGSGSGSGSALDIVMTQSPSLPVLTPGQSPASISCKXSNLLHSDGKTYLYVLQKP 175
 Qy 179 GQSPRLIYKVSNRDGVDPFRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTTFG 238

CC antibodies that can recognize the an amino acid moiety in human Mpl from
 CC positions 26-274; polynucleotides encoding these antibodies;
 CC polynucleotides hybridizable with the polynucleotides that encode the
 CC antibodies under stringent conditions and the antibodies encoded by which
 CC have equivalent activity as such antibodies; vectors containing these
 CC polynucleotides; host cells carrying such polynucleotides or vectors; and
 CC drug compositions containing these antibodies. The antibodies are
 CC applicable in drug compositions for use in chemotherapy of
 CC thrombocytopenia and other diseases associated with Mpl mutation such as
 CC congenital amegakaryocytic thrombocytopoenia. These antibodies can include
 CC anti-human Mpl antibody and variants like genetically-engineered anti-human
 CC Mpl diabody and (humanized) anti-human Mpl sc(Fv)2. Such antibodies are
 CC soluble to enhance binding ability. The diabody and sc(Fv)2 show high
 CC agonistic activity against anti-human Mpl antibody, such activity of which
 CC is comparable or superior to that of the natural ligand, human TPO. This
 CC is the amino acid sequence of a single chain antibody hVB22B q-wz5.
 XX
 SQ Sequence 524 AA;

Query Match 68.1%; Score 952.5; DB 9; Length 524;
 Best Local Similarity 74.8%; Pred. No. 3.8e-60;
 Matches 190; Conservative 25; Mismatches 30; Indels 9; Gaps 4;
 QY 2 AQVLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNTH 59
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 78
 19 SQVLVQSGPEVKKPGASVKVSCKASGYTFTNSWNVRQRPGLGLEWIGRIYPGGETI 78
 QY 60 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVVYCARDPRGDDDE---PYWQGGTLTV 116
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 135
 79 YNGKFRVVTITADESTSTAYMELSLRSDDTAVVYCA---RGYDDYSFAYWGGTITV 135
 QY 117 SSGGGSGGGSGGGSGGSEIVLTQSPSLPVLTPVTLPQASISCRSSQNLVYSGNTYLSWFQ 176
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 195
 136 SSGGGSGGGSGGGSGGSDIVMTQSPSLPVTTPGPASISCRSSKSLHSNGNTLYWFOQ 195
 QY 177 RPSQSPRLIYKVNRSQVDPDRPSGSGGTDFTLKISRVEAEDIGVYCMQGTTHPPRT 236
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 254
 196 KPGQAPRLIYRMSNLASGVDPDRPSGSGGTFTLKISRVEAEDVGYYCMQHIEY-PFT 254
 QY 237 FGQGTKEIKRGLG 250
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 268
 255 FGQGTKEIKGGG 268

RESULT 10
 AEE02143
 ID AEE02143 standard; protein; 524 AA.
 XX
 AC AEE02143;
 XX

26-JAN-2006 (first entry)
 DE Anti-Mpl sc(Fv)2 antibody hVB22B q-wz5.
 XX
 KW Antibody; myeloproliferative leukemia virus oncogene; Mpl;
 KW antibody engineering; humanized antibody; platelet; thrombocytopenia;
 KW hemostatic; hematological disease; ds; coding sequence;
 KW single chain antibody.
 XX
 OS Homo sapiens.
 OS Mus musculus.
 OS Synthetic.
 XX
 PN WO2005107784-A1.
 XX
 PD 17-NOV-2005.
 XX
 PF 11-MAY-2005; 2005WO-JP008592.
 XX
 PR 11-MAY-2004; 2004JP-00141618.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX

PI Tsuchiya M, Orita T, Matsumoto Y;
 XX WPI; 2005-779430/79.
 DR N-PSDB; AEE02145.
 XX
 PT Remedies for thrombopenia as well as agents for increasing number of
 PT platelets comprises e.g. Mpl-recognizing antibodies with high agonistic
 PT activity against anti-human Mpl antibody.
 XX
 PS Example 3; SEQ ID NO 101; 71pp; Japanese.
 XX
 CC The invention relates to remedies for thrombocytopenia comprising an
 CC sc(Fv)2 (a double single chain antibody comprising 2 heavy chains and 2
 CC light chains, often referred to as a diabody) with binding activity
 CC toward TPO (thrombopoietin) receptor (Mpl), myeloproliferative leukemia
 CC virus oncogene) as active ingredient. The antibodies comprise human
 CC framework regions (FR) grafted to mouse complementarity region (CDR).
 CC Also included are remedies for thrombocytopenia containing antibodies
 CC that can bind with human Mpl and monkey Mpl, remedies for
 CC thrombocytopenia containing antibodies with agonistic activity against
 CC Mpl, remedies for thrombocytopenia containing antibodies (which have a
 CC heavy-chain variable region with CDR1, CDR2 and CDR3 that have amino acid
 CC sequences of AEE02043-AEE02045, respectively, and/or a light-chain
 CC variable region with CDR1, CDR2 and CDR3 that have amino acid sequences
 CC of AEE02046-AEE02048, or antibodies having amino acid sequences based on
 CC AEE02043-AEE02048 for modification by substitution, deletion, addition
 CC and/or insertion of some amino acids), remedies for thrombocytopenia
 CC containing antibodies that can recognize amino acid sites of human Mpl
 CC from positions 126-274, and agents for increasing platelets containing
 CC the active ingredients as already specified for the thrombocytopenia
 CC remedies. The remedies are for the treatment of thrombocytopenia. Such
 CC antibodies are extremely effective because they have high agonistic
 CC activity against anti-human Mpl antibody, and the activity is comparable
 CC or superior to that of human TPO (thrombopoietin) which is a natural
 CC ligand. The present sequence represents a humanized single chain antibody
 CC (or diabody) derived from the mouse anti-Mpl antibody (VB22B). Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 524 AA;

Query Match 68.1%; Score 952.5; DB 9; Length 524;
 Best Local Similarity 74.8%; Pred. No. 3.8e-60;
 Matches 190; Conservative 25; Mismatches 30; Indels 9; Gaps 4;
 QY 2 AQVLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNTH 59
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 78
 19 SQVLVQSGPEVKKPGASVKVSCKASGYTFTNSWNVRQRPGLGLEWIGRIYPGGETI 78
 QY 60 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVVYCARDPRGDDDE---PYWQGGTLTV 116
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 135
 79 YNGKFRVVTITADESTSTAYMELSLRSDDTAVVYCA---RGYDDYSFAYWGGTITV 135
 QY 117 SSGGGSGGGSGGGSGGSEIVLTQSPSLPVLTPVTLPQASISCRSSQNLVYSGNTYLSWFQ 176
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 195
 136 SSGGGSGGGSGGGSGGSDIVMTQSPSLPVTTPGPASISCRSSKSLHSNGNTLYWFOQ 195
 QY 177 RPSQSPRLIYKVNRSQVDPDRPSGSGGTDFTLKISRVEAEDIGVYCMQGTTHPPRT 236
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 254
 196 KPGQAPRLIYRMSNLASGVDPDRPSGSGGTFTLKISRVEAEDVGYYCMQHIEY-PFT 254
 QY 237 FGQGTKEIKRGLG 250
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 268
 255 FGQGTKEIKGGG 268

RESULT 11
 AEC92009
 ID AEC92009 standard; protein; 247 AA.
 XX
 AC AEC92009;
 XX

The invention relates to remedies for thrombocytopenia comprising an sc(Fv)2 (a double single chain antibody comprising 2 heavy chains and 2 light chains, often referred to as a diabody) with binding activity toward TPO (thrombopoietin) receptor (Mpl, myeloproliferative leukemia virus oncogene) as active ingredient. The antibodies comprise human framework regions (FR) grafted to mouse complementarity region (CDR). Also included are remedies for thrombocytopenia containing antibodies that can bind with human Mpl and monkey Mpl, remedies for thrombocytopenia containing antibodies with agonistic activity against Mpl, remedies for thrombocytopenia containing antibodies (which have a heavy-chain variable region with CDR1, CDR2 and CDR3 that have amino acid sequences of AEE02043-AEE02045, respectively, and/or a light-chain variable region with CDR1, CDR2 and CDR3 that have amino acid sequences of AEE02046-AEE02048, or antibodies having amino acid sequences based on AEE02043-AEE02048 for modification by substitution, deletion, addition and/or insertion of some amino acids), remedies for thrombocytopenia containing antibodies that can recognize amino acid sites of human Mpl from positions 126-274, and agents for increasing platelets containing the active ingredients as already specified for the thrombocytopenia remedies. The remedies are for the treatment of thrombocytopenia. Such antibodies are extremely effective because they have high agonistic activity against anti-human Mpl antibody, and the activity is comparable or superior to that of human TPO (thrombopoietin) which is a natural ligand. The present sequence represents a humanized single chain antibody (or diabody) derived from the mouse anti-Mpl antibody (VR22B). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published](http://www.wipo.int/pub/published) pct sequences.

Antibody: myeloproliferative leukemia virus oncoqene: Mpl;

Db 136 SSGGGSGGGSGGGSDIVMTQSPSLUPVTPGEPASISCRSSKSLHNSGNTLYWFLQ 195
 Qy 177 RFGQSPRLIYKVNDRDGVDPDRFSGSGSGTDTLTKISRVEADIGVYICMQGTHWPPRT 236
 Db 196 KFGQSPQLLIYRMSNLASGVDPDRFSGSGSGTDTLTKISRVEADGVYICMQHIEY-PPT 254
 Qy 237 FGGQTKVIRKGLG 250
 Db 255 FGGQTKLEIKGGG 268

RESULT 14
 AEA43937
 ID AEA43937 standard; protein; 524 AA.
 XX AC AEA43937;
 XX DT 25-AUG-2005 (first entry)
 XX DE Anti-TPO single chain antibody hVB22B p-z.
 XX KW hemostatic; single chain antibody; protein purification;
 KW antibody engineering; antibody production; thrombocytopenia; hemostatic;
 KW hematological disease; hVB22B p-z.
 XX OS Homo sapiens.
 XX PN WO2005056604-A1.
 XX PD 23-JUN-2005.
 XX PF 10-DEC-2004; 2004WO-JP018506.
 XX PR 12-DEC-2003; 2003JP-00415746.
 XX PR 12-MAR-2004; 2004JP-00071763.
 XX PR 27-AUG-2004; 2004JP-00248323.
 XX PA (CHUS) CHUGAI SEIYAKU KK.
 XX PI Tsunoda H, Nakano K, Orita T, Tsuchiya M, Hirata Y;
 XX WPI; 2005-445149/45.
 XX DR N-PSDB; AEA43936.
 XX PT Anti-Mpl antibodies including antihuman Mpl antibody and variants like
 PT genetically-engineered antihuman Mpl diabody, applicable in drug
 PT compositions for use in chemotherapy of thrombocytopenia and other
 PT diseases associated with Mpl mutation.
 XX PS Claim 29; SEQ ID NO 2; 106pp; Japanese.
 CC The invention describes an antibody is a single-chain polypeptide with
 CC TPO (thrombopoietin) receptor (Mpl)-binding activity that contains the 2
 CC heavy-chain variable regions and 2 light-chain variable regions. Also
 CC described are: a chimeric antibody for binding with Mpl; an antibody for
 CC binding with soluble Mpl; antibodies for binding with human Mpl and
 CC monkey Mpl; antibodies with agonistic activity against human Mpl and
 CC monkey Mpl; antibodies with binding activity for soluble Mpl; antibodies
 CC with TPO agonistic activity; antibodies containing a heavy-chain variable
 CC region with CDR (complementary-determinant regions 1, 2 and 3);
 CC antibodies containing a light-chain variable region with CDR 1, 2 and 3;
 CC antibodies containing heavy and light-chain variable regions with CDR 1,
 CC 2 and 3 selected from the already-specified groups in 18 combinations;
 CC antibodies containing a heavy-chain variable region with FRL 1, 2, 3 and 4;
 CC antibodies containing a light-chain variable region FRL 1, 2, 3 and 4;
 CC epitope-recognizing antibodies that can recognize the antibodies with the
 CC heavy and light-chain variable regions and amino acid sequences;
 CC antibodies that can recognize the amino acid moiety in human Mpl from
 CC positions 26-274; polynucleotides encoding these antibodies;
 CC polynucleotides hybridizable with the polynucleotides that encode the
 CC antibodies under stringent conditions and the antibodies encoded by which
 CC have equivalent activity as such antibodies; vectors containing these

CC polynucleotides; host cells carrying such polynucleotides or vectors; and
 CC drug compositions containing these antibodies. The antibodies are
 CC applicable in drug compositions for use in chemotherapy of
 CC thrombocytopenia and other diseases associated with Mpl mutation such as
 CC congenital amegakaryocytic thrombocytopenia. These antibodies can include
 CC antihuman Mpl antibody and variants like genetically-engineered antihuman
 CC antihuman Mpl antibody and (humanized) antihuman Mpl sc(Fv)2. Such antibodies are
 CC soluble to enhance binding ability. The diabody and sc(Fv)2 show high
 CC agonistic activity against antihuman Mpl antibody, such activity of which
 CC is comparable or superior to that of the natural ligand, human TPO. This
 CC is the amino acid sequence of a single chain antibody hVB22B p-z.
 XX
 SQ Sequence 524 AA;
 Query Match 67.8%; Score 947.5; DB 9; Length 524;
 Best Local Similarity 74.8%; Pred. No. 8.6e-60;
 Matches 190; Conservative 24; Mismatches 31; Indels 9; Gaps 4;
 Qy 2 AQVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGITWVQAPQGLEWGWGI--NNGNTH 59
 Db 19 SQVQLVQSGPEVKKPKGASVKVSKASGYTFTNSWMNVRQRPCKGLEWGWGRIYPGDETI 78
 Qy 60 YAQKFGQGRVMTTDTSTSTAYMELSLRSDDTAVYVCARDPRGDDE--PYWQGGTLVTV 116
 Db 79 YNGKFRVTVITADESTSTAYMELSLRSEDATVYYCA--RGYDDYSFAYWQGGTIVV 135
 Qy 117 SSGGGSGGGSGGGSGSEIVLTQSPLSLPTVLGQPASISCRSSQNIVYSDGNTYLSWFQQ 176
 Db 136 SSGGGSGGGSGGGSGSDIVMTQSAISLPTVTPGEPASISCRSSKSLHNSGNTLYWFFQQ 195
 Qy 177 RFGQSPRLIYKVNDRDGVDPDRFSGSGSGTDTLTKISRVEADIGVYICMQGTHWPPRT 236
 Db 196 KFGQSPQLLIYRMSNLASGVDPDRFSGSGSGTDTLTKISRVEADGVYICMQHIEY-PPT 254
 Qy 237 FGGQTKVIRKGLG 250
 Db 255 FGGQTKLEIKGGG 268

RESULT 15
 AEE02103
 ID AEE02103 standard; protein; 524 AA.
 XX AC AEE02103;
 XX DT 26-JAN-2006 (first entry)
 XX DE Anti-Mpl sc(Fv)2 antibody hVB22B p-z.
 XX KW Antibody; myeloproliferative leukemia virus oncogene; Mpl;
 KW antibody engineering; humanized antibody; platelet; thrombocytopenia;
 KW hemostatic; hematological disease; ds; coding sequence;
 KW single chain antibody.
 XX OS Homo sapiens.
 XX OS Mus musculus.
 XX OS Synthetic.
 XX PN WO2005107784-A1.
 XX PF 17-NOV-2005.
 XX PF 11-MAY-2005; 2005WO-JP008592.
 XX PR 11-MAY-2004; 2004JP-00141618.
 XX PA (CHUS) CHUGAI SEIYAKU KK.
 XX PI Tsuchiya M, Orita T, Matsumoto Y;
 XX WPI; 2005-779430/79.
 XX DR N-PSDB; AEE02102.
 XX

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:14:57 ; Search time 50 Seconds
(without alignments)
456.910 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVQLVQSGAEVKKPGASV.....KVEIKRGLGGLVDYKDDDK 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC_Celerra_SID33/ptodata/2/iaa/5_COMB.pep.*
- 2: /EMC_Celerra_SID33/ptodata/2/iaa/6_COMB.pep.*
- 3: /EMC_Celerra_SID33/ptodata/2/iaa/7_COMB.pep.*
- 4: /EMC_Celerra_SID33/ptodata/2/iaa/H_COMB.pep.*
- 5: /EMC_Celerra_SID33/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /EMC_Celerra_SID33/ptodata/2/iaa/RE_COMB.pep.*
- 7: /EMC_Celerra_SID33/ptodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	914.5	65.4	535	2	US-08-983-035A-38
2	914.5	65.4	535	2	US-09-968-851A-38
3	909.5	65.1	243	2	US-09-297-181-2
4	905	64.7	267	2	US-09-419-788-30
5	883	63.2	353	2	US-09-203-958A-4
6	881.5	63.1	365	2	US-08-875-811-53
7	881.5	63.1	366	2	US-08-875-811-55
8	870	62.2	257	2	US-09-419-788-113
9	866	61.9	281	2	US-09-025-769B-178
10	866	61.9	281	2	US-09-490-070A-178
11	866	61.9	281	2	US-09-490-153-178
12	866	61.9	281	2	US-09-490-324-178
13	856	61.2	249	3	US-09-661-992B-84
14	850	60.8	249	2	US-09-726-219A-190
15	850	60.8	249	2	US-09-196-522-190
16	848	60.7	269	2	US-08-646-265A-109
17	828	59.2	247	2	US-09-227-693-34
18	828	59.2	248	1	US-08-331-397B-34
19	828	59.2	248	1	US-08-331-397B-34
20	828	59.2	248	1	US-08-759-804A-34
21	828	59.2	260	1	US-08-447-402-1
22	826	59.1	269	2	US-09-070-408-132
23	823	58.9	263	1	US-08-752-844-66
24	823	58.9	263	2	US-09-293-533-66
25	821	58.7	673	2	US-09-423-439-32
26	816	58.4	252	1	US-08-133-804-4

27	816	58.4	252	1	US-08-461-838-4	Sequence 4, Appli
28	816	58.4	252	1	US-08-461-886-4	Sequence 4, Appli
29	816	58.4	288	2	US-09-423-439-38	Sequence 38, Appl
30	808	57.8	246	1	US-08-257-341-7	Sequence 7, Appli
31	808	57.8	367	1	US-08-257-341-5	Sequence 5, Appli
32	807	57.7	289	2	US-09-184-658-63	Sequence 63, Appl
33	807	57.7	289	2	US-09-504-262D-63	Sequence 63, Appl
34	802	57.4	267	2	US-09-485-737B-2	Sequence 2, Appli
35	802	57.4	267	2	US-10-071-485-2	Sequence 2, Appli
36	794	56.8	711	2	US-09-485-737B-90	Sequence 90, Appl
37	794	56.8	711	2	US-10-071-485-90	Sequence 90, Appl
38	793	56.7	541	2	US-09-485-737B-85	Sequence 85, Appl
39	793	56.7	541	2	US-10-071-485-85	Sequence 85, Appl
40	790	56.5	240	1	US-08-488-113B-148	Sequence 148, App
41	790	56.5	240	1	US-08-477-484B-148	Sequence 148, App
42	790	56.5	240	1	US-08-646-360-148	Sequence 148, App
43	790	56.5	240	2	US-08-839-765-148	Sequence 148, App
44	790	56.5	240	2	US-09-136-389-148	Sequence 148, App
45	790	56.5	240	2	US-09-610-838-148	Sequence 148, App

ALIGNMENTS

RESULT 1

US-08-983-035A-38
; Sequence 38, Application US/08983035A
; Patent No. 6326464
; GENERAL INFORMATION:
; APPLICANT: CONSEILLER, EMMANUEL
; TITLE OF INVENTION: P53 PROTEIN VARIANTS AND THERAPEUTICAL
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS: USES THEREOF
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/983,035A
; FILING DATE: 20-Feb-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR96/01111
; FILING DATE: 17-JUL-1996
; APPLICATION NUMBER: FR 95/08729
; FILING DATE: 19-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Strause, William L.
; REGISTRATION NUMBER: 47,114
; REFERENCE/DOCKET NUMBER: 03804.0142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 535 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-08-983-035A-38

Query Match 65.4%; Score 914.5; DB 2; Length 535;

[illegible]

RESULT 2
 US-09-968-851A-38
 ; Sequence 38, Application US/09968851A
 ; Patent No. 6933373
 ; GENERAL INFORMATION:
 ; APPLICANT: CONSEILLER, EMMANUEL
 ; BRACCO, LAURENT
 ; TITLE OF INVENTION: P53 PROTEIN VARIANTS AND THERAPEUTICAL
 ; USES THEREOF
 ; NUMBER OF SEQUENCES: 59
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FINNEGAN, HENDERSON, PARABOW, GARRETT &
 ; DUNNER, LLP
 ; STREET: 1300 I Street, NW
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/968,851A
 ; FILING DATE: 03-Oct-2001
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/983,035
 ; FILING DATE: 20-FEB-1998
 ; APPLICATION NUMBER: PCT/FR96/01111
 ; FILING DATE: 17-JUL-1996
 ; APPLICATION NUMBER: FR 95/08729
 ; FILING DATE: 19-JUL-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Strauss, William L.
 ; REGISTRATION NUMBER: 47,114
 ; REFERENCE/DOCKET NUMBER: 03804.0142-01000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-408-4000
 ; TELEFAX: 202-408-4400
 ; INFORMATION FOR SEQ ID NO: 38:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 535 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 38:
 US-09-968-851A-38

Query Match	65.4%;	Score	914.5;	DB 2;	Length	535;			
Best Local Similarity	69.9%;	Pred. No.	3.7e-67;						
Matches	174;	Conservative	34;	Mismatches	36;	Indels	5;	Gaps	3;
Qy	1	MAQVQLVQSGAEYVKRPGASVKVSKASGVTFSNYGITVRRAPGQGLEWMGWIN--NGNT	58						
Db	1	MAQVQLQESGAELVSGSASVKLSCTASGFNIKDYMHVWKVQRPQEGLEWIGWDIPENGDT	60						
Qy	59	HYAKKQGRVTMTDTSTAYMELSLSDDTAVVYCARDPRGDDPEYKQGCTILTVSS	118						
Db	61	EYAPKPGKATMTADTSSNTAYLQLSSLASEDTAVVYC--NFYGDALDYVWGQGTITVTVSS	118						
Qy	119	GGSGSGGGSGGGGSRIVLTQSPFLSFVTLTGQPASISCRSSQNLVYSDGNTYILSWFQQRP	178						
Db	119	GGSGSGGGSGGGSDVLMTQTETLTLSVTIGQPASISCKSSQSLLSDSGKTYILNWLQRP	178						
Qy	179	QGSFRLIYVSNRDSGVDRPFGSGSGDTFTLKI SRVAEADIGVYVCWGTHWPPRTFG	238						
Db	179	QGSFRLIYLVKLDSGVDPDRFTFGSGSGDTFTLUKINRVEADIGVYVCWGTH--SPLTFG	237						
Qy	239	QGTKVEIKR	247						
Db	238	AGTKLEIKR	246						

```

RESULT 3
US-09-297-181-2
; Sequence 2, Application US/09297181
; Patent No. 6852509
; GENERAL INFORMATION:
; APPLICANT: Bracco, Laurent
; APPLICANT: Debussche, Laurent
; TITLE OF INVENTION: ANTI-PS3 SINGLE-CHAIN ANTIBODY FRAGMENTS AND THEIR USES
; FILE REFERENCE: ST96030-US
; CURRENT APPLICATION NUMBER: US/09/297,181
; CURRENT FILING DATE: 1999-04-26
; EARLIER APPLICATION NUMBER: PCT/FR97/01921
; EARLIER FILING DATE: 1997-10-27
; EARLIER APPLICATION NUMBER: FR96/13176
; EARLIER FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-297-181-2

```

Query Match	65.1%	Score	909.5;	DB 2;	Length	243;	
Best Local Similarity	70.3%;	Pred. No.	3.9e-67;				
Matches	173;	Conservative	33;	Mismatches	35;	Gaps	5;
Qy	3	QVQLVDSGAEVKPGASVKVSCKASGYTFSNYSGITWVRQAAPGGGLEWMGHIN--NGNTHY	60				
Db	1	QVQLQQSGAELVRSGASVKLSCTASGFNIKDYTMHWVKQRPEQGLEIWIWDPENGDTEY	60				
Qy	61	AQKFGQGVTTMTDTSTNTAYMEILSRSLSDTA VVYVCARDPRGDEDPYWGOGTLVTIVSSGG	120				
Db	61	APFKQKATMTADTSNTATAYLQLSSLASEDAVYYC--NFYGDALDWGOGTTVTIVSSGG	118				
Qy	121	GGSGGGGGGGGGSEIVLITQSPLSPVLTGOPASISCRSSONLYVSDGNTYLSWFQRPQG	180				
Db	119	GGSGGGGGGGGSDVLMQTPIFLSVTIGOPASISCKSSGLDSDGKTLYLNLLQRPGQ	178				
Qy	181	SPRLLIYKNRDSGVDPDRFSGSGGTDFTLKISRVEAEDIGVYCMQGTHTWPRTFGQG	240				
Db	179	SPKRLLIYNKLDGSVPDRFTGSGSGTDFTLKINRVEAEDLGIVYCWQGTG-SPLTFGAG	237				
Qy	241	TKVEIK	246				
Db	238	TKLEIK	243				

```

; PRIOR APPLICATION NUMBER: 60/067232
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic construct
US-09-203-958A-4

Query Match 63.2%; Score 883; DB 2; Length 353;
Best Local Similarity 66.1%; Pred. No. 8.9e-65;
Matches 168; Conservative 39; Mismatches 35; Indels 12; Gaps 4

QY 2 AQVQLVQSQAEEVKPGASVKVSKASGYTFSNYCITWVROAPGQGLEWMGWINN-----GN 57
DB 37 SEIQLQQTGPGLVKPGASVKISCKASGYSTFDIIFWVKSHGKSLWTG--NNPPYGS 94
QY 58 THYAQKFGQGRVTMTDTSTAYMELRSLRSDDTAVYYCAR-----DPRGDDEPYWGQGT 112
DB 95 TSNLAFKKGKATLVDKSSSTAYMQLNSLTSEDSAVYCVRGVYVYSSYEAFYWGQGT 154
QY 113 LVTVSSGGGGSGGGGGGGSEIVLQSPSLPVTILGQPASISCRSQNLVYSGNTYLS 172
DB 155 LVTVSAGGGGGGGGGGGGGSDVMTQTPLTSLITIGQPASISCKSSQSLDSDGKTYLN 214
QY 173 WFQORQGSPRRLLIYKVNEDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYVCMQGTHW 232
DB 215 WLLQRPQGSPTRLIYLVSKLDSGVPDRFSGSGSGTDFTLKISRVEAEDLGIYVCMQGAHF 274
QY 233 PPRTFGGQGTKEIK 246
DB 275 -PQTFGGGTKEIK 287

RESULT 6
US-08-875-811-53
; Sequence 53, Application US/08875811
; Patent No. 6045793
; GENERAL INFORMATION:
; APPLICANT: Rybak, Susanna M.
; APPLICANT: Newton, Dianne L.
; APPLICANT: Bogue, Lluís
; APPLICANT: Wlodawer, Alexander
; TITLE OF INVENTION: Recombinant Ribonuclease Proteins
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,811
; FILING DATE: 19-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/02588
; FILING DATE: 19-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/011,800
; FILING DATE: 21-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Paris, Susan K.
; REGISTRATION NUMBER: 41,739

```

; REFERENCE/DOCKET NUMBER: 015280-244100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-875-811-53

Query Match 63.1%; Score 881.5; DB 2; Length 365;
Best Local Similarity 67.3%; Pred. No. 1.2e-64;
Matches 167; Conservative 37; Mismatches 37; Indels 7; Gaps 4;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGIVTWVROAPGGLWGMWIN--NGNTHY 60
DB 119 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGIVTWVROAPGGLWGMWIN--NGNTHY 178
QY 61 AQKFGQRTVMTDTSTAYMELRSLRSDDTAVYYCARDP--RGDDEPYWGGTGLTVTVSSG 119
DB 179 ADDFKGRFAFSLTSAAYLIQINLNKNEDEATYFCARFAIKGD---YWGQGTITVTVSSG 235
QY 120 GGGSGGGSGGGSGGGSEIVLTQSPSLPVTLPGLQSPASISCRSSQNLVYSDGNTYLSWFOQRPG 179
DB 236 GGGSGGGSGGGSGGGSDIVLTQSPFNPVTLTGTSASISCRSTKSLHNSGITYLYWYLRKPG 295
QY 180 QSPRLIYKVSNRDGVDPDRSGSGSGTDFTLKISRVEAEIDGVYCMQGTHTWPPRTFGQ 239
DB 296 QSPQLLIYQMSNLASGVDPDRFSSSGSGTDFTLRISRVEAEIDGVYCAQNLE--IPRTFGG 354
QY 240 GTKVEIKR 247
DB 355 GTKLEIKR 362

RESULT 7
US-08-875-811-55
; Sequence 55, Application US/08875811
; Patent No. 6045793
; GENERAL INFORMATION:
; APPLICANT: Rybak, Susanna M.
; APPLICANT: Newton, Dianne L.
; APPLICANT: Boque, Lluis
; APPLICANT: Wlodawer, Alexander
; TITLE OF INVENTION: Recombinant Ribonuclease Proteins
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,811
; FILING DATE: 19-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/02588
; FILING DATE: 19-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/011,800
; FILING DATE: 21-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Paris, Susan K.
; REGISTRATION NUMBER: 41,739

; REFERENCE/DOCKET NUMBER: 015280-244100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-875-811-55

Query Match 63.1%; Score 881.5; DB 2; Length 366;
Best Local Similarity 67.3%; Pred. No. 1.2e-64;
Matches 167; Conservative 37; Mismatches 37; Indels 7; Gaps 4;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGIVTWVROAPGGLWGMWIN--NGNTHY 60
DB 2 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGIVTWVROAPGGLWGMWIN--NGNTHY 61
QY 61 AQKFGQRTVMTDTSTAYMELRSLRSDDTAVYYCARDP--RGDDEPYWGGTGLTVTVSSG 119
DB 62 ADDFKGRFAFSLTSAAYLIQINLNKNEDEATYFCARFAIKGD---YWGQGTITVTVSSG 118
QY 120 GGGSGGGSGGGSGGGSEIVLTQSPSLPVTLPGLQSPASISCRSSQNLVYSDGNTYLSWFOQRPG 179
DB 119 GGGSGGGSGGGSGGGSDIVLTQSPFNPVTLTGTSASISCRSTKSLHNSGITYLYWYLRKPG 178
QY 180 QSPRLIYKVSNRDGVDPDRSGSGSGTDFTLKISRVEAEIDGVYCMQGTHTWPPRTFGQ 239
DB 179 QSPQLLIYQMSNLASGVDPDRFSSSGSGTDFTLRISRVEAEIDGVYCAQNLE--IPRTFGG 237
QY 240 GTKVEIKR 247
DB 238 GTKLEIKR 245

RESULT 8
US-09-419-788-113
; Sequence 113, Application US/09419788
; Patent No. 6825325
; GENERAL INFORMATION:
; APPLICANT: FISCHER, Rainer
; APPLICANT: SCHILLBERG, Stefan
; APPLICANT: NAHRING, Jorg
; APPLICANT: SACK, Markus
; APPLICANT: MONECKE, Michael
; APPLICANT: LIAO, Yu-Cal
; APPLICANT: SPIEGEL, Holger
; APPLICANT: ZIMMERMAN, Sabine
; APPLICANT: EMANS, Neil
; TITLE OF INVENTION: Molecular Pathogenicide Mediated Plant Disease
; TITLE OF INVENTION: Resistance
; FILE REFERENCE: 0147-0189P
; CURRENT APPLICATION NUMBER: US/09/419,788
; CURRENT FILING DATE: 1999-10-18
; EARLIER APPLICATION NUMBER: 98 11 9630.6 EP
; EARLIER FILING DATE: 1998-10-16
; EARLIER APPLICATION NUMBER: 66/BOM/1998 INDIA
; EARLIER FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic, no
; OTHER INFORMATION: natural origin
US-09-419-788-113

Query Match 62.2%; Score 870; DB 2; Length 257;
Best Local Similarity 66.8%; Pred. No. 7.3e-64;

Matches 171; Conservative 33; Mismatches 36; Indels 16; Gaps 5;

QY 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVROAPGQGLEWMGWI--NNGNT 58
Db 1 MAREVQLQSGAEVKKPGASVKVSCKASGYTFTNVMHVKQTTPGQGLEWIGAIYPRNGDT 60
QY 59 HYAQKFGQRTMTTDTSTSTAYMELRLSLRSDTAVYYCARDPRGDDPVGQGTLLVTYS 118
Db 61 SYNQKFGKATLTADKSSSTAYMQLSSLTSDSAVYYCAR-----PDVWAGATLLTYS 114
QY 119 GGG---GSG---GGSGGGGSEIIVLTQSPLSLPTVTLGQPASISCRSSQNLVYSDGNTYL 171
Db 115 GAGPTSGSGKPGPEGSTKGAADPVLMTQAPUTLSVTIGQPASISCKSSQSLDGDGKTYL 174
QY 172 SWFOQRPGQSPRLIYKVSNRDSDGVPRDFSGSGSDTFTLKISRVEADIGVYICMQGTH 231
Db 175 NWLLQRPGQSPKRLIYLVSKLDGVPDRFTGSGSDTFTLKISRVEADIGVYICMQGTH 234
QY 232 WPPRTFGQGTKEIKR 247
Db 235 F-PHTFGGKLEIKR 249

RESULT 9
US-09-025-769B-178
; Sequence 178, Application US/09025769B
; Patent No. 6300064
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769B
; FILING DATE: 18-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-025-769B-178

Query Match 61.9%; Score 866; DB 2; Length 281;
Best Local Similarity 66.4%; Pred. No. 1.7e-63;
Matches 170; Conservative 36; Mismatches 36; Indels 14; Gaps 5;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVROAPGQGLEWMGWI--NNGNT 60
Db 26 EVQLVESGGGLVQPGGSLRLSCLASGFTFSSYAMSVWRQAPGKLEWVSAISGSGSTYY 85
QY 61 AQKFGQRTMTTDTSTSTAYMELRLSLRSDTAVYYCARDPRGDD---EPYWGQGTLLVT 116
Db 86 ADSVKGRTFTISRDNISKNTLYLQWNSLRADTAVYYCAR--WGSDGFYAMDYWGQGTLLVT 143
QY 117 S-----SGGGSGGGSGGGGSEIIVLTQSPLSLPTVTLGQPASISCRSSQNLVYSDGNTYL 171
Db 144 SAAGGSGGGSGGGSGGGGSDIVMTQSPLSLPTVTPGEPASISCRSSQSLHNSGNYL 203
QY 172 SWFOQRPGQSPRLIYKVSNRDSDGVPRDFSGSGSDTFTLKISRVEADIGVYICMQGTH 231
Db 204 DWLQKPGQSPQLLYLGNRASGVPRDFSGSGSDTFTLKISRVEADIGVYICMQGTH 263
QY 232 WPPRTFGQGTKEIKR 247
Db 264 TPP-TFGQGTKEIKR 278

RESULT 10
US-09-490-070A-178
; Sequence 178, Application US/09490070A
; Patent No. 6696248
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Colin G. Sandercock, Esq. c/o Heller Ehrman
; STREET: 1666 K Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/490,070A
; FILING DATE: 24-Jan-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Colin G. Sandercock, Esq.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 37629-0005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 912-2000
; TELEFAX: (202) 912-2020
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 178:
US-09-490-070A-178

Query Match 61.9%; Score 866; DB 2; Length 281;
Best Local Similarity 66.4%; Pred. No. 1.7e-63;
Matches 170; Conservative 36; Mismatches 36; Indels 14; Gaps 5;

Qy		3	QVQLVQSGLAEVKPGASVKVSCKASGTTFSNYGIVTWVRQAPEGQGLEWNGWI--NNGNTHY	60
Dd		26	EVQLVESGGGLVQPGGSLRLSCAASGTFSSYAMSVWRQAPGKLEWYSATSGSGSTVY	85
Qy		61	AQKFGQGVITMTDTSTAYMELSRSDRTAVYYCARDPRGDD---EPYWGQGTLTIV	116
Dd		86	ADSVKGFTISRDNKNTLIQMNSLRAEDTAVYYCAR--WGGDGFYAMDYWGQGTILTV	143
Qy		117	S-----SGGGSGGGSGGGSEIVLTSQSPISLPVTLTGPASPISCRSSQNLYSDGNITYL	171
Dd		144	SSAGGSGGGSGGGSGGGSGDSIDVMTSQSPISLPVPTPGEPASISCRSSQSLHSHSNYNYL	203
Qy		172	SWFQQRPGQPRRLIYKVSNRDSGVDPDFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHT	231
Dd		204	DWTLOKFGQSPQLLIYLGSNNASGVDPDRFSGSGSGTDFTLKISRVEAEDVGVIYCOQHYY	263
Qy		232	WPPTFGQGTKVEIKR	247
Dd		264	TTP--TFGQGTKVEIKR	278

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

Query Match	61.9%	Score 866;	DB 2;	Length 281;
Best Local Similarity	66.4%;	Pred. NO. 1.7e-63;		
Matches 170; Conservative	36;	Mismatches 36;	Indels 14;	Gaps 5;

Qy	3	QVQLVQSGAEVKPKGASVTKVSCKASGTYFSSYGLTWRQAPGQGLIEWGWI--NNGNTHY	60
Db	26	EVQLVESGGGLVQPGGSLRLASCAASGTFSSYAMSWVRQAPGKGLIEWSAISGSGGSTVY	85
Qy	61	AKQFGQVMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDD---EPYWGQGTLTVV	116
Db	86	ADSVKGRFTISRDNKNTLIYQMNSLRAEDTAVYYCAR--WGGDGFYAMDYWGQGTLTVV	143
Qy	117	S-----SGGGSGGGSGGGGSEIVLTQSPLSLPTLTGQPASISCRSSQNLYSDGNTYL	171
Db	144	SSAGGSGGGSGGGSGGGSGGSDIVMTQSPLSLPTPGEPASISCRSSQSLILHSNGYNIL	203
Qy	172	SWFOORPGQPRRLIYKVSNRDSCVPDRFSGSGSGTQFTLKISRVEADIYGVYCMQGTH	231
Db	204	DWYLQKPGQSPQLLIYLGSNRRASGVPDRFSGSGSGTQFTLKISRVEADVGVIYCOQHVT	263
Qy	232	WPPRTFGQGTKVEIKR	247
Db	264	TPP--TFGGGTKVEIKR	278

RESULT 12
 US-09-490-324-178
 ; Sequence 178, Application US/09490324
 ; Patent No. 6828422
 ; GENERAL INFORMATION:
 ; APPLICANT: Knappik, Achim
 ; Pack, Peter
 ; Ilag, Vic
 ; Ge, Liming
 ; Moroney, Simon
 ; Plueckthun, Andreas
 ; TITLE OF INVENTION: Protein/(Poly)peptide libraries
 ; NUMBER OF SEQUENCES: 373
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
 ; STREET: 1251 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10021
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/490.324
 ; FILING DATE: 24-Jan-2000
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/025,769
 ; FILING DATE: 18-FEB-1998
 ; APPLICATION NUMBER: EP 95 11 3021.0
 ; FILING DATE: 18-AUG-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: James F. Haley, Jr., Esq.
 ; REGISTRATION NUMBER: 27,794
 ; REFERENCE/DOCKET NUMBER: MORPHO/5
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 596-9000
 ; TELEFAX: (212) 596-9090
 ; INFORMATION FOR SEQ ID NO: 178:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 281 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 178:
 US-09-490-324-178

Query Match	61.9%;	Score 866;	DB 2;	Length 281;
Best Local Similarity	66.4%;	Pred. No. 1.7e-63;		

Matches	170;	Conservative	36;	Mismatches	36;	Indels	14;	Gaps	5
Qy	3	QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNTHY	60						
Db	26	EVQLVESGGGLVQPGGSLRLSCAASGFTFSYAMSWVRQAPGKLEWVSAISGGSGSTYY	85						
Qy	61	AQKFGQGRVTMTDTSTSTAYMELRLSRDDTAVVYICARDPRGDD----EPYWGQGTILTVV	116						
Db	86	ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVVYICAR--WGGDGFYAMDYWGQGTILTVV	143						
Qy	117	S-----SGGGSGGGSGGGGSEIVLTQSPSLPVLPTLTGQPASISCRSSQNLVYSDGNLYL	171						
Db	144	SSAGGGSGGGSGGGGGGGSDIVMTQSPSLPVTPEGPASISCRSSQSLHNSGNYL	203						
Qy	172	SWFQRRPQGPRLRIYKVSNRDGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYCMQGT	231						
Db	204	DWYLQKPGQSPQLLIYLGSRASGVDPDRFSGSGSGTDFTLKISRVEAEDGVYVYCCQHYT	263						
Qy	232	WPPRTFGQTKVEIKR	247						
Db	264	TPP-TFGQTKVEIKR	278						
RESULT 13									
US-09-661-992B-84									
; Sequence 84, Application US/09661992B									
; Patent No. 7033590									
; GENERAL INFORMATION:									
; APPLICANT: Scheifflinger, Friedrich									
; APPLICANT: Kerschbaumer, Randolph									
; APPLICANT: Falkner, Falko-Guenther									
; APPLICANT: Dörner, Friedrich									
; APPLICANT: Baxter Aktiengesellschaft									
; TITLE OF INVENTION: Factor IX/Factor IXa Activating Antibodies and Antibody									
; TITLE OF INVENTION: Derivatives									
; FILE REFERENCE: 20695C-005900US									
; CURRENT APPLICATION NUMBER: US/09/661,992B									
; CURRENT FILING DATE: 2000-09-14									
; PRIOR APPLICATION NUMBER: AT A157600									
; PRIOR FILING DATE: 1999-09-14									
; NUMBER OF SEQ ID NOS: 112									
; SOFTWARE: PatentIn ver. 2.1									
; SEQ ID NO 84									
; LENGTH: 249									
; TYPE: PRT									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Description of Artificial Sequence: scFv from									
; OTHER INFORMATION: hybridoma cell line 193/K2									
US-09-661-992B-84									
Query Match 61.2%; Score 856; DB 3; Length 249;									
Best Local Similarity 65.2%; Pred. No. 9.9e-63;									
Matches 163; Conservative 38; Mismatches 43; Indels 6; Gaps 3;									
Qy	3	QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWINGN--THY	60						
Db	1	EVQLVESGGGLVQPGGSLRLSCAASGFTFSYAMSWVRQTPKRELEWATISGGSGSTYY	60						
Qy	61	AQKFGQGRVTMTDTSTSTAYMELRLSRDDTAVVYICARD---PRGDDPEYWGQGTILTVTS	117						
Db	61	PDSVVRGRFTISRDNKNTLYLQMSLSKSEDTAMVYCTRDGCHGVGSGFDYWGQGTILTVS	120						
Qy	118	SGGGSGGGSGGGGSEIVLTQSPSLPVLPTLTGQPASISCRSSQNLVYSDGNLYLSWFOOR	177						
Db	121	SGGGSGGSRASGGGSGQIVLTQSPSLPVSIGDQASISCRSSQSIHNSGNTTYLEWYLOK	180						
Qy	178	PGOSPRLIYKVSNRDGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYVYCMQGTWHPPTF	237						
Db	181	PGOSPQLLIYKVSNRFGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYVYVYCPQSH--VPWTF	239						
Qy	238	GGGTKVEIKR	247						

Db 240 GGGTKLEIKR 249

RESULT 14

US-09-726-219A-190

; Sequence 190, Application US/09726219A

; Patent No. 6806079

; GENERAL INFORMATION:

; APPLICANT: Cambridge Antibody Technology

; APPLICANT: Cambridge Antibody Technology Limited

; APPLICANT: Medical Research Council

; APPLICANT: McCafferty, John

; APPLICANT: Pope, Anthony

; APPLICANT: Johnson, Kevin

; APPLICANT: Hoogenboom, Hendricus

; APPLICANT: Griffiths, Andrew

; APPLICANT: Jackson, Ronald

; APPLICANT: Holliger, Kasper

; APPLICANT: Marks, James

; APPLICANT: Clarkson, Timothy

; APPLICANT: Chiswell, David

; APPLICANT: Winter, Gregory

; APPLICANT: Bonert, Timothy

; TITLE OF INVENTION: Methods for Producing Members of Specific Binding Pairs

; FILE REFERENCE: 213839-00013

; CURRENT APPLICATION NUMBER: US/09/726,219A

; CURRENT FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: GB 9015198.6

; PRIOR FILING DATE: 1990-07-10

; PRIOR APPLICATION NUMBER: GB 9022845.3

; PRIOR FILING DATE: 1990-10-19

; PRIOR APPLICATION NUMBER: GB 9022845.3

; PRIOR FILING DATE: 1990-10-19

; PRIOR APPLICATION NUMBER: GB 9024503.6

; PRIOR FILING DATE: 1990-11-12

; PRIOR APPLICATION NUMBER: GB 9104744.9

; PRIOR FILING DATE: 1991-03-06

; PRIOR APPLICATION NUMBER: GB 9110549.4

; PRIOR FILING DATE: 1991-05-15

; PRIOR APPLICATION NUMBER: PCT/GB91/01134

; PRIOR FILING DATE: 1991-07-10

; PRIOR APPLICATION NUMBER: US 07/971,857

; PRIOR FILING DATE: 1993-01-08

; PRIOR APPLICATION NUMBER: US 08/484,893

; PRIOR FILING DATE: 1995-06-07

; NUMBER OF SEQ ID NOS: 272

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 190

; LENGTH: 249

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: scFv form of the anti-oxazalone antibody NQ11

US-09-726-219A-190

Query Match 60.8%; Score 850; DB 2; Length 249;

Best Local Similarity 64.8%; Pred. No. 3.1e-62;

Matches 164; Conservative. 35; Mismatches 42; Indels 12; Gaps 4

Qy 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSSNYGTTWVRQAPGQGLEWGWING---NT 58

Db 1 QVQLQESGGGLVQPGGSLRLSCATSGFTFSNYYMGWVRQPPGKALEWLGSRNKNVGYTT 60

Qy 59 HYAQKFGQRTVMTTDTSTSTAYMELRSRSDTAVYTCARDPRGDE---PYWQSGTLV 114

Db 61 EYSASVGRFTISRDNFQSLILYLQINTLRDTSATYYCA---RGVDYGANFAYWQGGTLV 117

Qy 115 TVSSGGGGSGGGSGGGGSEIVLTQSPLSLPVTLGQPASISCRSSQNLYVSDGNTYLSWF 174

Db 118 TVSSGGGGSGGGSGGGGSDIELTQTPLSLPVSLGDAQSISCRSSQSVHNSNGNTYLEWY 177

Qy 175 QQRPGQSPRLIYKVSNRDSCVDPDRFGSGSGTDTFLKISRVEADIGVYICMQGTHHP 234

Db 178 LQKPGQSPKLLIYKVSNRFSGVDPDRFSGSGGTFTLTKISRVEAEDLGVYYCFQGSHP 236

QY 235 RTFGQGTKEIKR 247

Db 237 YTFGGGTTKLEIKR 249

RESULT 15

US-09-196-522-190
; Sequence 190, Application US/09196522
; Patent No. 6916605
; GENERAL INFORMATION:
; APPLICANT: Cambridge Antibody Technology
; APPLICANT: Cambridge Antibody Technology Limited
; APPLICANT: Medical Research Council
; APPLICANT: McCafferty, John
; APPLICANT: Pope, Anthony
; APPLICANT: Johnson, Kevin
; APPLICANT: Hoogenboom, Hendricus
; APPLICANT: Griffiths, Andrew
; APPLICANT: Jackson, Ronald
; APPLICANT: Holliger, Kasper
; APPLICANT: Marks, James
; APPLICANT: Clackson, Timothy
; APPLICANT: Chiswell, David
; APPLICANT: Winter, Gregory
; APPLICANT: Bonert, Timothy
; TITLE OF INVENTION: Methods for Producing Members of Specific Binding Pairs
; FILE REFERENCE: 213839-00004
; CURRENT APPLICATION NUMBER: US/09/196,522
; CURRENT FILING DATE: 1998-11-28
; PRIOR APPLICATION NUMBER: GB 9015198.6
; PRIOR FILING DATE: 1990-07-10
; PRIOR APPLICATION NUMBER: GB 9022845.3
; PRIOR FILING DATE: 1990-10-19
; PRIOR APPLICATION NUMBER: GB 9022845.3
; PRIOR FILING DATE: 1990-10-19
; PRIOR APPLICATION NUMBER: GB 9024503.6
; PRIOR FILING DATE: 1990-11-12
; PRIOR APPLICATION NUMBER: GB 9104744.9
; PRIOR FILING DATE: 1991-03-06
; PRIOR APPLICATION NUMBER: GB 9110549.4
; PRIOR FILING DATE: 1991-05-15
; PRIOR APPLICATION NUMBER: PCT/GB91/01134
; PRIOR FILING DATE: 1991-07-10
; PRIOR APPLICATION NUMBER: US 07/971,857
; PRIOR FILING DATE: 1993-01-08
; PRIOR APPLICATION NUMBER: US 08/484,893
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 272
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 190
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: scFv form of the anti-oxazolone antibody NQ11
US-09-196-522-190

Query Match 60.8%; Score 850; DB 2; Length 249;
Best Local Similarity 64.8%; Pred. No. 3.1e-62;
Matches 164; Conservative 35; Mismatches 42; Indels 12; Gaps 4;
QY 3 QVQLVQSGAEVKPKGASVKSCASGYTFNYSYGTWVRQAPGGLEWGWING-----NT 58
Db 1 QVQLQESGGGLVQPGGSLRLSCTSGFTFSNYMGWVRQPPGKALEWLGSRNKNVGYTT 60
QY 59 HYAQKFGQVRVTMTDTSTSTAYMELRSLRSDTAVYYCARDPRGDDE---PYWGQGTLV 114
Db 61 EYASVAGRTIISRDNFQSIYLQINTLRDSDSATYCA---RGYDYCAWFAYWGQGTLV 117
QY 115 TVSSGGGSGGGSGGGSGGSEIVLTQSPLSPLVTLTGQPASISCRSSQNLVYSDGNTYLSWF 174

Db 118 TVSSGGGSGGGSGGGSGGSDIELTQTPLSLPVSILGDQASISCRSSQSIVHSNGNTYLEWY 177
QY 175 QORPGQSPRRLIYKVSNRDSGVDPDRFSGSGGTFTLTKISRVEAEDIGVYYCMQGTHWPP 234
Db 178 LQKPGQSPKLLIYKVSNRFSGVDPDRFSGSGGTFTLTKISRVEAEDLGVYYCFQGSHP 236
QY 235 RTFGQGTKEIKR 247
Db 237 YTFGGGTTKLEIKR 249

Search completed: August 9, 2006, 13:16:20
Job time : 52 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:26:47 ; Search time 179 Seconds

(without alignments)
675,414 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVQLVQSGAEVKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1398	100.0	261	6	US-11-056-825-2
2	1382	98.9	259	6	US-11-056-825-7
3	1135	81.2	253	3	US-09-880-748-1964
4	1135	81.2	253	4	US-10-293-418-1964
5	1135	81.2	253	6	US-11-054-515-1964
6	1135	81.2	253	6	US-11-266-444-1964
7	963	68.9	249	6	US-11-056-825-4
8	952	68.1	247	6	US-11-056-825-8
9	933.5	66.8	274	4	US-10-257-864A-92
10	933.5	66.8	274	5	US-10-645-085A-92
11	929.5	66.5	274	4	US-10-138-505-26
12	929.5	66.5	274	4	US-10-257-864A-90
13	929.5	66.5	274	4	US-10-399-518-119
14	929.5	66.5	274	5	US-10-399-585-118
15	929.5	66.5	274	5	US-10-645-085A-90
16	928.5	66.4	274	4	US-10-138-505-32
17	928.5	66.4	274	4	US-10-221-131-97
18	928.5	66.4	274	4	US-10-399-518-121
19	928.5	66.4	274	5	US-10-399-585-120
20	924.5	66.1	244	3	US-09-880-748-1991
21	924.5	66.1	244	4	US-10-293-418-1991
22	924.5	66.1	244	6	US-11-054-515-1991
23	924.5	66.1	244	6	US-11-266-444-1991
24	921.5	65.9	271	4	US-10-138-505-30
25	921.5	65.9	271	4	US-10-257-864A-91
26	921.5	65.9	271	4	US-10-221-131-96
27	921.5	65.9	271	4	US-10-399-518-120

28	921.5	65.9	271	5	US-10-399-585-119	Sequence 119, App
29	921.5	65.9	271	5	US-10-645-085A-91	Sequence 91, Appl
30	920.5	65.8	271	4	US-10-138-505-34	Sequence 34, Appl
31	920.5	65.8	271	4	US-10-257-864A-93	Sequence 93, Appl
32	920.5	65.8	271	4	US-10-221-131-98	Sequence 98, Appl
33	920.5	65.8	271	4	US-10-399-518-122	Sequence 122, App
34	920.5	65.8	271	5	US-10-399-585-121	Sequence 121, App
35	920.5	65.8	271	5	US-10-645-085A-93	Sequence 93, Appl
36	920.5	65.8	271	4	US-10-257-864A-96	Sequence 96, Appl
37	920.5	65.8	533	4	US-10-221-131-101	Sequence 101, App
38	920.5	65.8	533	4	US-10-399-518-125	Sequence 125, App
39	920.5	65.8	533	5	US-10-399-585-124	Sequence 124, App
40	920.5	65.8	533	5	US-10-645-085A-96	Sequence 96, Appl
41	919.5	65.8	507	4	US-10-239-656-47	Sequence 47, Appl
42	919.5	65.8	510	4	US-10-239-656-48	Sequence 48, Appl
43	919.5	65.8	510	4	US-10-239-656-49	Sequence 49, Appl
44	919.5	65.8	532	6	US-11-036-098-18	Sequence 18, Appl
45	914.5	65.4	535	3	US-09-968-851-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1

US-11-056-825-2

; Sequence 2, Application US/11056825

; Publication NO. US20050255109A1

; GENERAL INFORMATION:

; APPLICANT: Feilding-Habermann, Brunhilde

; APPLICANT: Janda, Kim D.

; APPLICANT: Saven, Alan

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS

; FILE REFERENCE: SGRP-0042

; CURRENT APPLICATION NUMBER: US/11/056,825

; CURRENT FILING DATE: 2005-02-11

; PRIOR APPLICATION NUMBER: US 60/626,726

; PRIOR FILING DATE: 2004-11-10

; PRIOR APPLICATION NUMBER: US 60/544,807

; PRIOR FILING DATE: 2004-02-13

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 2

; LENGTH: 261

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: Synthetic Construct

US-11-056-825-2

Query Match 100.0%; Score 1398; DB 6; Length 261;
Best Local Similarity 100.0%; Pred. No. 2, 7e-93;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAQVQLVQSGAEVKPGASVVKSCASGVTFSNYGTVTRQAPQGGLWGWINGNTHY	60
Db	1	MAQVQLVQSGAEVKPGASVVKSCASGVTFSNYGTVTRQAPQGGLWGWINGNTHY	60
QY	61	AQKFGQRTVTTDTSTAYMELRSDDTAVYTCARDPRGDDEPYWGQGLTVTVSSGG	120
Db	61	AQKFGQRTVTTDTSTAYMELRSDDTAVYTCARDPRGDDEPYWGQGLTVTVSSGG	120
QY	121	GGSGGGGGGGGGSEIVLTQSPLSPTLQGPASISCRSSQNLVYSDGNTYLSWFOQRPQG	180
Db	121	GGSGGGGGGGGGSEIVLTQSPLSPTLQGPASISCRSSQNLVYSDGNTYLSWFOQRPQG	180
QY	181	SPRRLIYKVNRSQGVDPFRFGSGSGTDTFLTKISRVEAEDIGVYCMQGTHWPPRTFGQ	240
Db	181	SPRRLIYKVNRSQGVDPFRFGSGSGTDTFLTKISRVEAEDIGVYCMQGTHWPPRTFGQ	240
QY	241	TKVEIKRGLGLVDYKDDDK 261	
Db	241	TKVEIKRGLGLVDYKDDDK 261	


```
RESULT 2
US-11-056-825-7
; Sequence 7, Application US/11056825
; Publication No. US2005025109A1
; GENERAL INFORMATION:
; APPLICANT: Felsing-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Saven, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SCRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: US 60/544,807
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-7

Query Match          98.9%; Score 1382; DB 6; Length 259;
Best Local Similarity 99.6%; Pred. No. 3.8e-92;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWINNGNTHY 60
Db 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWINNGNTHY 60

Qy 61 AOKFQGRVTMTDTSTSTAYMELSLRSDDTAVVYCARDPRGDDDPYWGQGLTVTVSSGG 120
Db 61 AOKFQGRVTMTDTSTSTAYMELSLRSDDTAVVYCARDPRGDDDPYWGQGLTVTVSSGG 120

Qy 121 GSGGGGGGGGSEILVTSPLSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFOORPQG 180
Db 121 GSGGGGGGGGSEILVTSPLSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFOORPQG 180

Qy 181 SPRLLYKVSNRDGVPRDFSGSGSGTDFTLKISRVEADIGVYCMQGTWHPPTFCQG 240
Db 181 SPRLLYKVSNRDGVPRDFSGSGSGTDFTLKISRVEREDIGVYCMQGTWHPPTFCQG 240

Qy 241 TKVEIKRGLGLVDYKDD 259
Db 241 TKVEIKRGLGLVDYKDD 259

RESULT 3
US-09-880-748-1964
; Sequence 1964, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1964

Query Match          81.2%; Score 1135; DB 3; Length 253;
Best Local Similarity 86.2%; Pred. No. 2.8e-74;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKEPQASVKVSCKASGYTFTSISWVRQAPGQGLEWMGWSAYNGNTNY 60

Qy 61 AOKFQGRVTMTDTSTSTAYMELSLRSDDTAVVYCARDPRGDD-----PYWGQGLTVT 115
Db 61 AQLQGRVTMTDTSTSTAYMELSLRSDDTAVVYCARPNPYVDSSEGGPFDYWGQGLTMT 120

Qy 116 VSSGGGGGGGGGGG--EIVLTQSPSLPLPVTILGQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGGGGGGGGGG--EIVLTQSPSLPLPVTILGQPASISCRSSQSLVYSDGNTYLSW 180

Qy 174 FOORPQSPRLLYKVSNRDGVPRDFSGSGSGTDFTLKISRVEADIGVYCMQGTWHP 233
Db 181 FOORPQSPRLLYKVSNRDGVPRDFSGSGSGTDFTLKISRVEADVGVIYCMQATRW- 239

Qy 234 PRTFGQGTKEIKR 247
Db 240 PRTFGQGTKEIKR 253
```

```
RESULT 4
US-10-293-418-1964
; Sequence 1964, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1964

Query Match          81.2%; Score 1135; DB 4; Length 253;
Best Local Similarity 86.2%; Pred. No. 2.8e-74;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKEPQASVKVSCKASGYTFTSISWVRQAPGQGLEWMGWSAYNGNTNY 60
```

```

Db 181 FQRPQGSPRLIYKVNRSNDGVPDRFSGSGSGTFTTLKISRVEAEDGVVYCMQATRW - 239
Qy 234 PRTFQGQTKVEIKR 247
Db 240 PFTFGQGTQMEIKR 253

RESULT 6
US-11-266-444-1964
; Sequence 1964, Application US/11266444
; Publication No. US20060062789A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte
; FILE REFERENCE: PF523P1D1
; CURRENT APPLICATION NUMBER: US/11/266,444
; CURRENT FILING DATE: 2005-11-04
; PRIOR APPLICATION NUMBER: 09/880,746
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-266-444-1964

Query Match 81.2%; Score 1135; DB 6; Length 253;
Best Local Similarity 86.2%; Pred. No. 2.8e-74;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4

Qy 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGIMRWQAQPGGLEWGWGIN--NGNTHY 60
Db 1 QVQLQQSGAEVKKPGASVKVSCKASGYTFTSYGISWVRQAQPGGLEWGWISAYNGNTNY 60
Qy 61 AQKQGRVTMTDTSTSTAYMELSLRSDDTAVVYCARDPRGDDE----PYWGQGLT 115
Db 61 AQKQGRVTMTDTSTSTAYMELSLRSDDTAVVYCARNPYYDSSEGFDDYWGQGTMT 120
Qy 116 VSSGGSGSGGGSGGGGGS--BIVLTQPSLSPVTLGPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGSGSGGGSGGGGGSALDVMTQPSLSPVTLGPASISCRSSQSLVYSDGNTYLNW 180
Qy 174 FQRPQGSPRLIYKVNRSNDGVPDRFSGSGSGTFTTLKISRVEAEDGVVYCMQGTHTP 233
Db 181 FQRPQGSPRLIYKVNRSNDGVPDRFSGSGSGTFTTLKISRVEAEDGVVYCMQATRW - 239
Qy 234 PRTFQGQTKVEIKR 247
Db 240 PFTFGQGTQMEIKR 253

RESULT 7
US-11-056-825-4
; Sequence 4, Application US/11056825
; Publication No. US20050255109A1
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Saven, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SCR-0042
; CURRENT APPLICATION NUMBER: US/11/056,825

```

```
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 4
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-4

Query Match      68.9%; Score 963; DB 6; Length 249;
Best Local Similarity 72.0%; Pred. No. 7,7e-62;
Matches 190; Conservative 29; Mismatches 27; Indels 18; Gaps 5;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWIN--NGNT 58
Db 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWINPDSGDT 60

Qy 59 HYAQKFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSS 118
Db 61 NSAQFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSS 120

Qy 119 GGGSGGGSGGGSGGSEIVLTQSPSLPVTLGQPASISCRSSQNLVYSDGNTYLSWFOQRP 178
Db 61 NSAQFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSS 120

Qy 121 G-----GGGSGSEIVLTQSPGTLSPGERATLSCRASQSV-----SSSYLAWYQOKP 167
Db 121 G-----GGGSGSEIVLTQSPGTLSPGERATLSCRASQSV-----SSSYLAWYQOKP 167

Qy 179 GQSPRLIYKVNRSRGVDPFRFSGSGGTDFTLKISRVEAEDIGVYYCMQ-GTHWPPRTF 237
Db 168 GQAPRLIYGASSRATGIPDRFSGSGGTDFTLKISRLEPEDFAVYYCQYGS--SPRTF 225

Qy 238 GQGTKVEIKRGLGLVDYKDDDK 261
Db 226 GQGTKVDIKRGLGLVDYKDDDK 249

RESULT 8
US-11-056-825-8
; Sequence 8, Application US/11056825
; Publication No. US20050255109A1
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SCRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 8
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-8

Query Match      68.1%; Score 952; DB 6; Length 247;
Best Local Similarity 71.8%; Pred. No. 4,8e-61;
Matches 188; Conservative 29; Mismatches 27; Indels 18; Gaps 5;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWIN--NGNT 58
Db 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWINPDSGDT 60
```

```
Qy 59 HYAQKFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSS 118
Db 61 NSAQFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSS 120

Qy 119 GGGSGGGSGGGSGGSEIVLTQSPSLPVTLGQPASISCRSSQNLVYSDGNTYLSWFOQRP 178
Db 121 G-----GGGSGSEIVLTQSPGTLSPGERATLSCRASQSV-----SSSYLAWYQOKP 167

Qy 179 GQSPRLIYKVNRSRGVDPFRFSGSGGTDFTLKISRVEAEDIGVYYCMQ-GTHWPPRTF 237
Db 168 GQAPRLIYGASSRATGIPDRFSGSGGTDFTLKISRLEPEDFAVYYCQYGS--SPRTF 225

Qy 238 GQGTKVEIKRGLGLVDYKDDDK 259
Db 226 GQGTKVDIKRGLGLVDYKDDDK 247
```

RESULT 9

```
US-10-257-864A-92
; Sequence 92, Application US/10257864A
; Publication No. US20040058393A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, NAOSHI
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OH-EDA, MASAYOSHI
; APPLICANT: UNO, SHINSUKE
; APPLICANT: KIKUCHI, YASUFUMI
; APPLICANT: GOTEMBA-SHI, OHTOMO
; TITLE OF INVENTION: AGONIST ANTIODIES
; FILE REFERENCE: 065678-0107
; CURRENT APPLICATION NUMBER: US/10/257,864A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: PCT/JP01/03288
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/JP01/01912
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: JP2000-115246
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: JP2000-321821
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: JP2000-321822
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 92
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: amino acid sequence encoded by SEQ ID NO: 24
US-10-257-864A-92
```

```
Query Match      66.8%; Score 933.5; DB 4; Length 274;
Best Local Similarity 69.8%; Pred. No. 1.2e-59;
Matches 185; Conservative 31; Mismatches 34; Indels 15; Gaps 5;
```

```
Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNT 58
Db 21 MAQVQLQSGAEVKKPGASVKVSCKASGYTFANFVHHVKKPKPGQGLEWIGYIYPYNDGT 80

Qy 59 HYAQKFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEY--WGQGLTVTV 116
Db 81 KNEKPKRATLTSDKSTTAYMDLSSLASSEDGAVYYCA---RGYYTYDDWGQGLTVTV 137

Qy 117 SSGGGSGGGSGGGSGGSEIVLTQSPSLPVTLGQPASISCRSSQNLVYSDGNTYLSWFOQ 176
Db 138 SSGGGSGGGSGGGSGGSDVVMVTQSPSLPVLGDAQSISCRSSQSLVHNSGKTYLHWYIQ 197

Qy 177 RQGSPPRLIYKVNRSRGVDPFRFSGSGGTDFTLKISRVEAEDIGVYYCMQGTHWPPRT 236
Db 198 KPGQSPKLLIYKVNRSRGVDPFRFSGSGGTDFTLKISRVEAEDLGVYFCQSSTH-VPYT 256
```

```
QY 237 FGGTKVEIKRGLGLVDYKDDDK 261
Db 257 FGGTKLEIK-----DYKDDDK 274

RESULT 10
US-10-645-085A-92
; Sequence 92, Application US/10645085A
; Publication No. US20040258684A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, NAOSHI
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OH-EDA, MASAYOSHI
; APPLICANT: UNO, SHINSUKE
; APPLICANT: KIKUCHI, YASUFUMI
; APPLICANT: OHTOMO, TOSHIHIKO
; TITLE OF INVENTION: AGONIST ANTIBODIES
; FILE REFERENCE: 065678-0108
; CURRENT APPLICATION NUMBER: US/10/645,085A
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: PCT/JP01/03288
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/JP01/01912
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: JP 2000-115246
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: JP 2000-321821
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: JP 2000-321822
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 92
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: amino acid sequence encoded by SEQ ID NO: 24
US-10-645-085A-92

Query Match 66.8%; Score 933.5; DB 5; Length 274;
Best Local Similarity 69.8%; Pred. No. 1.2e-59;
Matches 185; Conservative 31; Mismatches 34; Indels 15; Gaps 5;

QY 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGITWVRQAPGQGLGWMGWI--NNGNT 58
Db 21 MAQVQLQSGAEIVKPGASVKMSCKASGYTFANVHWVKQKPGQGLEWIGYIYPNDGT 80

QY 59 HYAQKFGQGRVTMTDTSTAYMELSLRSDDTAVYYCARDPRGDDEPY--WGQGTLLTV 116
Db 81 KYNEKFKDKATLTSDKSTTAYMDLSLASEDSAVYYCA---RGYYTYDDWGQGTLLTV 137

QY 117 SSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 176
Db 138 SSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 197

QY 177 RFGQSPRLIYKVNRSQGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHMPRT 236
Db 198 KPGQSPKLLIYKVNRSQGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHMPRT 256

QY 237 FGGTKVEIKRGLGLVDYKDDDK 261
Db 257 FGGTKLEIK-----DYKDDDK 274

RESULT 11
US-10-138-505-26
; Sequence 26, Application US/10138505
; Publication No. US20030108546A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, Naoshi
; APPLICANT: UNO, Shinsuke
; APPLICANT: OH-EDA, Masayoshi
```

```
; APPLICANT: KIKUCHI, Yasufumi
; TITLE OF INVENTION: APOPTOSIS-INDUCING SINGLE-CHAIN FV
; FILE REFERENCE: 065678/0102
; CURRENT APPLICATION NUMBER: US/10/138,505
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US/09/523,095
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: JP 11-63557
; PRIOR FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-138-505-26

Query Match 66.5%; Score 929.5; DB 4; Length 274;
Best Local Similarity 69.1%; Pred. No. 2.2e-59;
Matches 183; Conservative 32; Mismatches 35; Indels 15; Gaps 5;

QY 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGITWVRQAPGQGLGWMGWI--NNGNT 58
Db 21 MAQVQLQSGGPDIVKPGASVKMSCKASGYTFVNVHVMHWKQKPGQGLEWIGYIYPNDGT 80

QY 59 HYAQKFGQGRVTMTDTSTAYMELSLRSDDTAVYYCARDPRGDDEPY--WGQGTLLTV 116
Db 81 KYNEKFKGKATLTSEKSSAAAYMELSLASEDSAVYYCA---RGYYYSYDDWGQGTLLTV 137

QY 117 SSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 176
Db 138 SSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 197

QY 177 RFGQSPRLIYKVNRSQGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHMPRT 236
Db 198 KPGQSPKLLIYKVNRSQGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHMPRT 256

QY 237 FGGTKVEIKRGLGLVDYKDDDK 261
Db 257 SGGGTKLEIK-----DYKDDDK 274

RESULT 12
US-10-257-864A-90
; Sequence 90, Application US/10257864A
; Publication No. US20040058393A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, NAOSHI
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OH-EDA, MASAYOSHI
; APPLICANT: UNO, SHINSUKE
; APPLICANT: KIKUCHI, YASUFUMI
; APPLICANT: GOTEMBA-SHI, OHTOMO
; TITLE OF INVENTION: AGONIST ANTIBODIES
; FILE REFERENCE: 065678-0107
; CURRENT APPLICATION NUMBER: US/10/257,864A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: PCT/JP01/03288
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/JP01/01912
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: JP2000-115246
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: JP2000-321821
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: JP2000-321822
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Mus musculus
```

FEATURE:
OTHER INFORMATION: amino acid sequence encoded by SEQ ID NO: 20
US-10-257-864A-90

Query Match 66.5%; Score 929.5; DB 4; Length 274;
Best Local Similarity 69.1%; Pred. No. 2.2e-59;
Matches 183; Conservative 32; Mismatches 35; Indels 15; Gaps 5;
QY 1 MAQVQLVQSGAEVKPKGASVKVCKASGYTFNNGITVWRQAPGQGLEWMGWI--NNGNT 58
DB 21 MAQVQLQSGDPLVKPGASVKMSCKASGYTFNVHVMHWKQKPGQGLEWIGIYIPYNDGT 80
QY 59 HYAQFQGRVTMTDTSTAYMELSLRSDTAVYYCARDPRGDEPY--WGQGLTVTV 116
DB 81 KYNEKFKGKATLTSEKSSAAVMELSLASEDSAVYYCA---RGYYSYDDWGQGLTVTV 137
QY 117 SSGGGSGGGSGGGSGGSIIVLTQSPLSIPVTLQGPASISCRSSQNLVYSDGNTYLSWFOQ 176
DB 138 SSGGGSGGGSGGGSGGSDVMTQTPLSLPVSLGDAQSISCRSSQLLHSGKNTYLQWYIQ 197
QY 177 RFGQSPRLIYKVSNRDGVDPDRFSGSGSGDTFTLKISRVEAEDIGVYCMQGTWHPPT 236
DB 198 KPGQSPKLLIYKVSNRFGVDPDRFSGSGSGDTFTLKISRVEAEDLGIVYFCSQSTH-VPYT 256
QY 237 FCGQTKVEIKRGLGLVDYKDDDDK 261
DB 257 SGGGTKLEIK-----DYKDDDDK 274

RESULT 13

US-10-399-518-119
Sequence 119, Application US/10399518
Publication No. US20040091475A1
GENERAL INFORMATION:
APPLICANT: TSUCHIYA, MASAYUKI
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: YABUTA, NAOHICO
APPLICANT: TSUNODA, HIROYUKI
APPLICANT: ORITA, TETSURO
TITLE OF INVENTION: DEGRADED TPO AGONIST ANTIBODY
FILE REFERENCE: 065678/0111
CURRENT APPLICATION NUMBER: US/10/399,518
CURRENT FILING DATE: 2003-09-25
PRIOR APPLICATION NUMBER: PCT/JP01/03288
PRIOR FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: JP 2001-277314
PRIOR FILING DATE: 2001-09-12
PRIOR APPLICATION NUMBER: JP 2000-321821
PRIOR FILING DATE: 2000-10-20
NUMBER OF SEQ ID NOS: 183
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 119
LENGTH: 274
TYPE: PRT
ORGANISM: Mus sp.
US-10-399-518-119

Query Match 66.5%; Score 929.5; DB 4; Length 274;
Best Local Similarity 69.1%; Pred. No. 2.2e-59;
Matches 183; Conservative 32; Mismatches 35; Indels 15; Gaps 5;
QY 1 MAQVQLVQSGAEVKPKGASVKVCKASGYTFNNGITVWRQAPGQGLEWMGWI--NNGNT 58
DB 21 MAQVQLQSGDPLVKPGASVKMSCKASGYTFNVHVMHWKQKPGQGLEWIGIYIPYNDGT 80
QY 59 HYAQFQGRVTMTDTSTAYMELSLRSDTAVYYCARDPRGDEPY--WGQGLTVTV 116
DB 81 KYNEKFKGKATLTSEKSSAAVMELSLASEDSAVYYCA---RGYYSYDDWGQGLTVTV 137
QY 117 SSGGGSGGGSGGGSGGSIIVLTQSPLSIPVTLQGPASISCRSSQNLVYSDGNTYLSWFOQ 176
DB 138 SSGGGSGGGSGGGSGGSDVMTQTPLSLPVSLGDAQSISCRSSQLLHSGKNTYLQWYIQ 197

QY 177 RFGQSPRLIYKVSNRDGVDPDRFSGSGSGDTFTLKISRVEAEDIGVYCMQGTWHPPT 236
DB 198 KPGQSPKLLIYKVSNRFGVDPDRFSGSGSGDTFTLKISRVEAEDLGIVYFCSQSTH-VPYT 256
QY 237 FCGQTKVEIKRGLGLVDYKDDDDK 261
DB 257 SGGGTKLEIK-----DYKDDDDK 274

RESULT 14

US-10-399-585-118
Sequence 118, Application US/10399585
Publication No. US20040242847A1
GENERAL INFORMATION:
APPLICANT: TSUCHIYA, MASAYUKI
APPLICANT: UNO, SHINSUKE
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: YABUTA, NAOHICO
APPLICANT: TSUNODA, HIROYUKI
TITLE OF INVENTION: DEGRADED AGONIST ANTIBODY
FILE REFERENCE: 065678/0112
CURRENT APPLICATION NUMBER: US/10/399,585
CURRENT FILING DATE: 2003-04-18
PRIOR APPLICATION NUMBER: PCT/JP01/09260
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: PCT/JP01/03288
PRIOR FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: PCT/JP01/01912
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: JP 2001-277314
PRIOR FILING DATE: 2001-09-12
PRIOR APPLICATION NUMBER: JP 2000-321821
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: JP 2000-321822
PRIOR FILING DATE: 2000-10-20
NUMBER OF SEQ ID NOS: 177
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 118
LENGTH: 274
TYPE: PRT
ORGANISM: Mus sp.
US-10-399-585-118

Query Match 66.5%; Score 929.5; DB 5; Length 274;
Best Local Similarity 69.1%; Pred. No. 2.2e-59;
Matches 183; Conservative 32; Mismatches 35; Indels 15; Gaps 5;
QY 1 MAQVQLVQSGAEVKPKGASVKVCKASGYTFNNGITVWRQAPGQGLEWMGWI--NNGNT 58
DB 21 MAQVQLQSGDPLVKPGASVKMSCKASGYTFNVHVMHWKQKPGQGLEWIGIYIPYNDGT 80
QY 59 HYAQFQGRVTMTDTSTAYMELSLRSDTAVYYCARDPRGDEPY--WGQGLTVTV 116
DB 81 KYNEKFKGKATLTSEKSSAAVMELSLASEDSAVYYCA---RGYYSYDDWGQGLTVTV 137
QY 117 SSGGGSGGGSGGGSGGSIIVLTQSPLSIPVTLQGPASISCRSSQNLVYSDGNTYLSWFOQ 176
DB 138 SSGGGSGGGSGGGSGGSDVMTQTPLSLPVSLGDAQSISCRSSQLLHSGKNTYLQWYIQ 197
QY 177 RFGQSPRLIYKVSNRDGVDPDRFSGSGSGDTFTLKISRVEAEDIGVYCMQGTWHPPT 236
DB 198 KPGQSPKLLIYKVSNRFGVDPDRFSGSGSGDTFTLKISRVEAEDLGIVYFCSQSTH-VPYT 256
QY 237 FCGQTKVEIKRGLGLVDYKDDDDK 261
DB 257 SGGGTKLEIK-----DYKDDDDK 274

RESULT 15

US-10-645-085A-90
Sequence 90, Application US/10645085A
Publication No. US20040258684A1

GENERAL INFORMATION:
APPLICANT: FUKUSHIMA, NAOSHI
APPLICANT: TSUCHIYA, MASAYUKI
APPLICANT: OH-EDA, MASAYOSHI
APPLICANT: UNO, SHINSUKE
APPLICANT: KIKUCHI, YASUFUMI
APPLICANT: OHTOMO, TOSHIHIKO
TITLE OF INVENTION: AGONIST ANTIBODIES
FILE REFERENCE: 065678-0108
CURRENT APPLICATION NUMBER: US/10/645,085A
CURRENT FILING DATE: 2002-10-07
PRIOR APPLICATION NUMBER: PCT/JP01/03288
PRIOR FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: PCT/JP01/01912
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: JP 2000-115246
PRIOR FILING DATE: 2000-04-17
PRIOR APPLICATION NUMBER: JP 2000-321821
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: JP 2000-321822
PRIOR FILING DATE: 2000-10-20
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 90
LENGTH: 274
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
OTHER INFORMATION: amino acid sequence encoded by SEQ ID NO: 20
US-10-645-085A-90

Query Match 66.5%; Score 929.5; DB 5; Length 274;
Best Local Similarity 69.1%; Pred. No. 2.2e-59;
Matches 183; Conservative 32; Mismatches 35; Indels 15; Gaps 5;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMCHI--NNGNT 58
Db 21 MAQVQLQSGPDLVKPGASVVMCKASGYTFVNHVMHWVKQKPGQGLEWIGYIYPNDGT 80
Qy 59 HYAOKFQGRVTMTTDTSTAYMELRSLSDDTAVYYCARDPRGDDEPY--WGQGTLLTV 116
Db 81 KYNEKFKGKATLTSEKSSAAYMELSSLSAEDSAVYICA---RGYYSYDDWGQGTLLTV 137
Qy 117 SSGGGSGGGSGGGSGGGGIVLTQSPLSLPLVTLGPASISCRSSQNLVYSDGNTYLSWFQQ 176
Db 138 SSGGGSGGGSGGGSGGGSDVMTQTPLSLPVSLGDAQASISCRSSQSLHSGKNTYLQWYLQ 197
Qy 177 RQGSPRLIYKVNRSRGVDPDFSGSGGTDFTLKISRVEAEDIGVYVCMQGTTHWPRT 236
Db 198 KPGQSPKLLIYKVNRSRGVDPDFSGSGGTDFTLKISRVEAEDIGVYVCSQSTH-VPYPT 256
Qy 237 FQGQTKVEIKRGLGLVDYKDDDDK 261
Db 257 SGGGTKLEIK-----DYKDDDDK 274

Search completed: August 9, 2006, 13:30:14
Job time : 180 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:27:22 ; Search time 33 Seconds
(without alignments)
532.377 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVQLVQSGAEVKKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 239914 seqs, 67312017 residues

Total number of hits satisfying chosen parameters: 239914

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	943	67.5	256	7	US-11-337-300-278
2	863	61.7	252	7	US-11-337-300-290
3	846	60.5	697	7	US-11-155-444-2
4	846	60.5	697	7	US-11-156-109-2
5	846	60.5	731	7	US-11-155-444-6
6	846	60.5	731	7	US-11-156-109-6
7	813	58.2	252	6	US-10-684-237-4
8	810	57.9	247	7	US-11-337-300-65
9	799.5	57.2	244	6	US-10-539-402-32
10	791	56.6	254	7	US-11-337-300-216
11	786	56.2	254	7	US-11-337-300-284
12	783.5	56.0	255	7	US-11-337-300-242
13	783	56.0	254	7	US-11-337-300-218
14	782	55.9	254	7	US-11-337-300-276
15	780	55.8	254	7	US-11-337-300-274
16	779	55.7	254	7	US-11-337-300-240
17	777	55.6	247	7	US-11-337-300-36
18	774	55.4	254	7	US-11-337-300-214
19	774	55.4	269	6	US-10-539-402-1
20	772.5	55.3	253	7	US-11-337-300-244
21	768	54.9	252	7	US-11-337-300-264
22	768	54.9	252	7	US-11-337-300-288
23	767	54.9	256	7	US-11-337-300-286
24	765	54.7	252	7	US-11-337-300-232
25	765	54.7	252	7	US-11-337-300-270

26	765	54.7	252	7	US-11-337-300-272	Sequence 272, Appl
27	764.5	54.7	242	7	US-11-183-325-20	Sequence 20, Appl
28	764.5	54.7	255	7	US-11-337-300-238	Sequence 238, Appl
29	764	54.6	252	7	US-11-337-300-252	Sequence 252, Appl
30	763.5	54.6	279	6	US-10-968-757-2	Sequence 2, Appl
31	763	54.6	252	7	US-11-337-300-266	Sequence 266, Appl
32	762.5	54.5	242	7	US-11-183-325-18	Sequence 18, Appl
33	761	54.4	252	7	US-11-337-300-228	Sequence 228, Appl
34	761	54.4	252	7	US-11-337-300-280	Sequence 280, Appl
35	759.5	54.3	248	6	US-10-539-402-17	Sequence 17, Appl
36	759	54.3	252	7	US-11-337-300-226	Sequence 226, Appl
37	758.5	54.3	254	7	US-11-075-891-6	Sequence 6, Appl
38	758	54.2	252	7	US-11-337-300-256	Sequence 256, Appl
39	757	54.1	252	7	US-11-337-300-254	Sequence 254, Appl
40	756	54.1	252	7	US-11-337-300-230	Sequence 230, Appl
41	755	54.0	260	6	US-10-991-309B-160	Sequence 160, Appl
42	754.5	54.0	254	7	US-11-075-891-8	Sequence 8, Appl
43	754	53.9	252	7	US-11-337-300-268	Sequence 268, Appl
44	753.5	53.9	251	7	US-11-337-300-212	Sequence 212, Appl
45	753	53.9	249	7	US-11-337-300-69	Sequence 69, Appl

ALIGNMENTS

RESULT 1

US-11-337-300-278
; Sequence 278, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruif, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudsmit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT APPLICATION NUMBER: US/11/337,300
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 278
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: SC03-053
US-11-337-300-278

Query Match 67.5%; Score 943; DB 7; Length 256;
Best Local Similarity 74.2%; Pred. No. 2.1e-62;
Matches 190; Conservative 19; Mismatches 35; Indels 12; Gaps 5;
QY 1 MAQVQLVQSGAEVKKPGASVKVSCKASGTFSTNYGIVTWVROAPQGLEWVGWIN--NGNT 58
Db 2 MAQVQLVQSGAEVKKPGASVKVSCKASGTFSTNYAMHVVRRQPGQLRWGNADKQGT 61
QY 59 HYAQKFGQRTVTTDTSTAYMELSLRSDDTAVVYCARDP-----RGDDEPYWGQT 112
Db 62 KYQKFGQRTVTTDTSTAYMELSLRSDDTAVVYCARDP-----RGDDEPYWGQT 119
QY 113 LVTVSSGGSGGGSGGGSGGGGGS-EIVLTQSPFLSLPVTLPGLQPASISCRSSQNLVYSDGNTYL 171
Db 120 TVTVSSGGSGGGSGGGSGGGSGGGGSLRSDDTAVVYCARDP-----RGDDEPYWGQT 179
QY 172 SNFQRRPGSPRLIYKVSNRSDGVPDRFGSGSGGTDFTLKISRVEAEDTIGVYCMQGT 231
Db 180 DWYLOKPGSPQLIYKVSNRSDGVPDRFGSGSGGTDFTLKISRVEAEDTIGVYCMQGT 239
QY 232 WPPRTFGQGTKEIKR 247
Db 240 -TPLTFGQGTLEIKR 254

```
RESULT 2
US-11-337-300-290
; Sequence 290, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Cruce Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruif, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudmit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT APPLICATION NUMBER: US/11/337,300
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 290
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: SC03-059
US-11-337-300-290

Query Match 61.7%; Score 863; DB 7; Length 252;
Best Local Similarity 68.6%; Pred. No. 1.6e-56;
Matches 175; Conservative 26; Mismatches 40; Indels 14; Gaps 6;

Qy 1 MAQVLVQSGAEVKKPGASVKVSCKAS-GYTSNVTGIVTWRQAPQGGLWGMWGN--NGN 57
Db 2 MAQVLVQSGAEVKKPGASVKVSCKASGGGVPRNAINVWRQAPQGGLWGMWGNPSGGS 61

Qy 58 THYAKQFGKRVMTTDTSTSTAYMELRSRSDTAVYVCARDPRGDDP---PYWQGGTL 113
Db 62 TSYAKQFGKRVLTDTSTSTAYMELSLRSRSDTAVYVCARDPRGTRGRYMDVWKGKTT 121

Qy 114 VTVSSGGSGGGSGGGSGGGS-BIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLS 172
Db 122 VTVSSGTGGSGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 176

Qy 173 WFQKPGQSPRLIYKVNDRSGVDPFRSGSGGTDFTLKISRVAEDIGVYCYWGQTHW 232
Db 177 WYQKPGQAPRLIYDASNRATGIPARFSGSGGTDFTLTISLPEPDPAVYCYCQQRSNW 236

Qy 233 PPRFGQGTQVEIKR 247
Db 237 PP-AFGGQTKVEIKR 250

RESULT 3
US-11-155-444-2
; Sequence 2, Application US/11155444
; Publication No. US20060104971A1
; GENERAL INFORMATION:
; APPLICANT: GARBER, ELLEN
; APPLICANT: BAILLY, VERONIQUE
; APPLICANT: BROWNING, JEFFREY L.
; TITLE OF INVENTION: MULTIVALENT LYMPHOTOXIN BETA RECEPTOR AGONISTS AND
; FILE REFERENCE: BGN168CN
; CURRENT APPLICATION NUMBER: US/11/155,444
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/US03/041393
; PRIOR FILING DATE: 2003-12-22
; PRIOR APPLICATION NUMBER: 60/435,154
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/435,185
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 2
; LENGTH: 697

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Heavy chain of
; OTHER INFORMATION: hCBell/hbHA10 bispecific-1 antibody construct
US-11-155-444-2

Query Match 60.5%; Score 846; DB 7; Length 697;
Best Local Similarity 67.2%; Pred. No. 7.7e-55;
Matches 166; Conservative 28; Mismatches 43; Indels 10; Gaps 4;

Qy 2 AQVQLVQSGAEVKKPGASVKVSCKASGYTFSNVTGIVTWRQAPQGGLWGMWGN-- 59
Db 459 SQVQLVQSGAEVKKPGSSVKVSCKASGYTFTTYYLHWVRQAPQGGLWGMWIPGNVHAQ 518

Qy 60 YAKQFGKRVMTTDTSTSTAYMELRSRSDTAVYVCARDPRGDDP---PYWQGGTTLVTSSG 119
Db 519 YNEKFGKRVITADKSTSTAYMELSLRSRSDTAVYVCARSWEG--FPYWGQGTTLVTSSG 576

Qy 120 GGGSGGGSGGGSGGGSEIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFOORPG 179
Db 577 GGGSGGGSGGGSGGSDIQMTQSPSSLSASVGDRTVITCKASQNV-----GINVAWYQKPKG 631

Qy 180 QSPRLIYKVNDRSGVDPFRSGSGGTDFTLKISRVAEDIGVYCYWGQTHHPPRTFGQ 239
Db 632 KAPKSLISSASYRYSGVPSRFSGSGGTDFTLTISLPEDFATYFCQYDTY--PFTFGQ 690

Qy 240 GTKVEIK 246
Db 691 GTKVEIK 697

RESULT 4
US-11-156-109-2
; Sequence 2, Application US/11156109
; Publication No. US20060134102A1
; GENERAL INFORMATION:
; APPLICANT: GILL, ALAN
; APPLICANT: LEPAGE, DOREEN
; TITLE OF INVENTION: LYMPHOTOXIN BETA RECEPTOR AGENTS IN COMBINATION WITH
; FILE REFERENCE: BGN171CN
; CURRENT APPLICATION NUMBER: US/11/156,109
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/US03/041243
; PRIOR FILING DATE: 2003-12-22
; PRIOR APPLICATION NUMBER: 60/435,185
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/435,154
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 2
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Heavy chain of
; OTHER INFORMATION: hCBell/hbHA10 bispecific-1 antibody construct
US-11-156-109-2

Query Match 60.5%; Score 846; DB 7; Length 697;
Best Local Similarity 67.2%; Pred. No. 7.7e-55;
Matches 166; Conservative 28; Mismatches 43; Indels 10; Gaps 4;

Qy 2 AQVQLVQSGAEVKKPGASVKVSCKASGYTFSNVTGIVTWRQAPQGGLWGMWGN-- 59
Db 459 SQVQLVQSGAEVKKPGSSVKVSCKASGYTFTTYYLHWVRQAPQGGLWGMWIPGNVHAQ 518

Qy 60 YAKQFGKRVMTTDTSTSTAYMELRSRSDTAVYVCARDPRGDDP---PYWQGGTTLVTSSG 119
Db 519 YNEKFGKRVITADKSTSTAYMELSLRSRSDTAVYVCARSWEG--FPYWGQGTTLVTSSG 576
```



```
; OTHER INFORMATION: 26-10 sFv'
US-10-684-237-4

Query Match      58.2%; Score 813; DB 6; Length 252;
Best Local Similarity 63.5%; Pred. No. 7.5e-53;
Matches 160; Conservative 40; Mismatches 46; Indels 6; Gaps 4;

Qy 3 QVQLVQSGAEVKKPKQASVKVSKASGYTFSNYGITWVRQAPQGQLEWGMWIN--NGNTHY 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 EVQLQQSGPELVKPGASVRMSCKSSGYITFDYFNMVVRQSHGSLDYIGYISPYSGVTGY 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 61 AOKFQGRVTMTDTSTAYMELSLRSDDTAVYYCARDPRGD-DEPYWGQGLTVTVSSG 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 NQKFKGKATLVTKSSSTAYMELSLTSSEDAVYVYVYVYVYVYVYVYVYVYVYVYVSS- 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 120 GGGSGGGSGGGGSSIVLTQSPVLTLPVLTGQAPASISCRSSQNLVYSDGNTYLSWFOQRP 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 -SGSSSSGSSGSSGVDMVTQTPSLPVLGQASISCRSSQSLVHSNGNTYLNWYVQKAG 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 180 QSPRLIYKVNRRDGVDPDRFSGSGGTDTFTLKISRVEAEDIGVYVYCMQGTHTWPPRTFQ 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 180 QSPKLLIYKVNRRFSGVDPDRFSGSGGTDTFTLKISRVAEDLGIYFCSTQTHVPP-TEGG 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 240 GTKVEIKRGLGG 251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 GTKLEIKRSGGG 250
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-11-337-300-65
; Sequence 65, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruif, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT APPLICATION NUMBER: US/11/337,300
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: SC03-012
US-11-337-300-65

Query Match      57.9%; Score 810; DB 7; Length 247;
Best Local Similarity 66.5%; Pred. No. 1.2e-52;
Matches 167; Conservative 20; Mismatches 48; Indels 16; Gaps 6;

Qy 1 MAQVLVQSGAEVKKPKGASVKVSKASGYTFSNYGITWVRQAPQGQLEWGMWIN--NGNT 58
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 MAQVLVQSGAEVKKPKGASVKVSKASGYTFSNYGITWVRQAPQGQLEWGMWISVNGNT 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 59 HYAQFQGRVTMTDTSTAYMELSLRSDDTAVYYCARDPRGDDEPYWGQGLTVTVSS 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 NYAQKLGRTVTMTDTSTAYMELSLRSDDTAVYYCARDPRKSPFDSMGQGLTVTVSR 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 119 GGGSGGGSGGGGSGSEIVLTQSPVLTLPVLTGQAPASISCRSSQNLVYSDGNTYLSWFOQRP 178
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 GGGSGGGSGGGGSGS-ELTQDP-AVSVALGQTVRITCGDSLRSY-----YASMYQKRP 174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 179 GQSPRLIYKVNRRDGVDPDRFSGSGGTDTFTLKISRVEAEDIGVYVYCMQGTHTWPP 234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 GQAPVLVYKGNRRFSGVDPDRFSGSGGNTASLTITGAQAEADYVYCNRRDSSGNH--- 231
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 235 RTFGQGTKEI 245
```

```
Db 232 VVFGGTKLTV 242
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-10-539-402-32
; Sequence 32, Application US/10539402
; Publication No. US20060115477A1
; GENERAL INFORMATION:
; APPLICANT: Xerion Pharmaceuticals AG
; APPLICANT: Tufts University
; TITLE OF INVENTION: Neuropilin-1 Inhibitor
; FILE REFERENCE: XE12EP
; CURRENT APPLICATION NUMBER: US/10/539,402
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US 60/435,893
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: EP 03000615
; PRIOR FILING DATE: 2003-01-15
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 244
; TYPE: PRT
; ORGANISM: human
; OTHER INFORMATION:
US-10-539-402-32

Query Match      57.2%; Score 799.5; DB 6; Length 244;
Best Local Similarity 61.6%; Pred. No. 7.2e-52;
Matches 154; Conservative 36; Mismatches 37; Indels 23; Gaps 5;

Qy 12 EVKFGASVKVSKASGYTFSNYGITWVRQAPQGQLEWGMWIN--NGNTHYAKFQGRVT 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 EVKFGASVKVSKASGYTFSNYGITWVRQAPQORLEWGMWINAGNGNTKYSKPFQGRVT 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 70 MTTDTSTAYMELSLRSDDTAVYYCARD-----PRGDDPYWGQGLTVTVSS 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 LTGDTSTSTAYMELSLTSSEDAVYYCARDQVYFSGSYIRPSFD---FWGRGTLTVTVSS 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 118 GGGSGGGSGGGGSGSEIVLTQSPVLTLPVLTGQAPASISCRSSQNLVYSDGNTYLSWFOQR 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 GGGSGGGSGGGGSGGSDIQTQSPVLTSLASIGDRVTITCRASEGIYH-----WLAWYQOK 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 178 PQQSPRLIYKVNRRDGVDPDRFSGSGGTDTFTLKISRVEAEDIGVYVYCMQGTHTWPPRTF 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 PKGAPKLLIYKASSLASGAPSRFSGSGGTDTFTLTSSLPQDDPFATYVYVYVYVYVYVYV 231
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 238 GQGTKEIKR 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 GGGTKLEIKR 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-11-337-300-216
; Sequence 216, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruif, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudsmit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT APPLICATION NUMBER: US/11/337,300
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 216
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
```

```
; OTHER INFORMATION: SC03-021
US-11-337-300-216

Query Match      56.6%; Score 791; DB 7; Length 254;
Best Local Similarity 63.2%; Pred. No. 3.1e-51;
Matches 160; Conservative 33; Mismatches 52; Indels 8; Gaps 6;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWINNNGN--T 58
Db 2 MAQVQLVQSGTEVKKPGESLISKSGSGYFITWIGWVRQMPGKGLWMMGIYFGDSET 61
Qy 59 HYAQKFGQRTVMTTDTSTSTAYMELRLSRSDTAVYVCARDPRGDDEP--YWGQGTLLTV 116
Db 62 RYSPFQGVQVITISADKSLNTAYLQWSSLSKASDTAIYCA--GGSGISTPMDVMWQGTITTV 120
Qy 117 SSGGGGGGGGGGGGG--EIVLTQSPSLPLPTLQGPASISCRSSQNLVYSDGN--TYLSWF 174
Db 121 SSGTGGSGGTGGTGGTDDVMTQSPDSLAVSLGERATINCKSSQSVLHSSNNKNYLAWY 180
Qy 175 QORPGQSPRLIYKVNRRDGVPRFSGSGGTDTFTLKISRVEAEDIGVYCMQGTTHWPP 234
Db 181 QOKPGQPPKLLIYWASTRQSGVPRFSGSGGTDTFTLTISLQAEADVAVYICQYVSTPP 240
Qy 235 RTFGQGTKEIKR 247
Db 241 -TFGQGTKEIKR 252

RESULT 11
US-11-337-300-284
; Sequence 284, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruif, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudsmit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 284
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURES:
; OTHER INFORMATION: SC03-056
US-11-337-300-284

Query Match      56.2%; Score 786; DB 7; Length 254;
Best Local Similarity 62.5%; Pred. No. 7.3e-51;
Matches 158; Conservative 36; Mismatches 51; Indels 8; Gaps 6;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWINNNGN--T 58
Db 2 MAEVQLVQSGTEVKKPGESLISKSGSGYFITWIGWVRQMPGKGLWMMGIYFGDSET 61
Qy 59 HYAQKFGQRTVMTTDTSTSTAYMELRLSRSDTAVYVCARDPRGDDEP--YWGQGTLLTV 116
Db 62 RYSPFQGVQVITISADKSLNTAYLQWSSLSKASDTAIYCA--GGSGISTPMDVMWQGTITTV 120
Qy 117 SSGGGGGGGGGGGGG--EIVLTQSPSLPLPTLQGPASISCRSSQNLVYSDGN--TYLSWF 174
Db 121 SSGTGGSGGTGGTGGTDDVMTQSPDSLAVSLGERATINCKSSQSVLHSSNNKNYLAWY 180
Qy 175 QORPGQSPRLIYKVNRRDGVPRFSGSGGTDTFTLKISRVEAEDIGVYCMQGTTHWPP 234
Db 181 QOKPGQPPKLLIYWASTRQSGVPRFSGSGGTDTFTLTISLQAEADVAVYICQYVSTPP 239
Qy 235 RTFGQGTKEIKR 247

; OTHER INFORMATION: SC03-056
US-11-337-300-284

Query Match      56.0%; Score 783.5; DB 7; Length 255;
Best Local Similarity 62.7%; Pred. No. 1.1e-50;
Matches 158; Conservative 34; Mismatches 53; Indels 7; Gaps 5;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWINNNGN--T 58
Db 2 MAEVQLVQSGAEVKKPGESLISKSGSGYFITWIGWVRQMPGKGLWMMGIYFGDSET 61
Qy 59 HYAQKFGQRTVMTTDTSTSTAYMELRLSRSDTAVYVCARDPRGDDEP--YWGQGTLLTV 116
Db 62 RYSPFQGVQVITISADKSLNTAYLQWSSLSKASDTAIYCA--GGSGISTPMDVMWQGTITTV 120
Qy 117 SSGGGGGGGGGGGGG--EIVLTQSPSLPLPTLQGPASISCRSSQNLVYSDGN--TYLSWF 174
Db 121 SSGTGGSGGTGGTGGTDDIQTQSPDSLAVSLGERATINCKSSQSVLHSSNNKNYLAWY 180
Qy 175 QORPGQSPRLIYKVNRRDGVPRFSGSGGTDTFTLKISRVEAEDIGVYCMQGTTHWPP 234
Db 181 QOKPGQPPKLLIYWASTRQSGVPRFSGSGGTDTFTLTISLQAEADVAVYICQYVSTPL 240
Qy 235 RTFGQGTKEIKR 246
Db 241 YSFGQGTKEIKR 252

RESULT 13
US-11-337-300-218
; Sequence 218, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruif, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudsmit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 218
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial sequence
US-11-337-300-242
; Sequence 242, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruif, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudsmit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 242
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURES:
; OTHER INFORMATION: SC03-035
US-11-337-300-242
```

; FEATURE:
; OTHER INFORMATION: SC03-022
US-11-337-300-218

Query Match 56.0%; Score 783; DB 7; Length 254;
Best Local Similarity 63.2%; Pred. No. 1.2e-50;

Matches 160; Conservative 33; Mismatches 52; Indels 8; Gaps 6;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGITHVRQAPGQGLEWMGWINNEN--T 58

Db 2 MAQVQLVQSGAEVKKPGESLKISKCKSGYGYFTYWGVRQMPGKGLWVGIIYPGDSSET 61

Qy 59 HYAQKFGQRTVMTTDTSTSTAYMELSLRSDDTAVYICARDPRGDDP--YWGQGTLLTV 116

Db 62 RYSPFQGVVTSADKSIINTAYLQWSSLKASDTAIYCA--GSGISTPMDVWGQGTLLTV 120

Qy 117 SSGGGSGGGSGGGGGS--EIVLTQSPSLPVTLPQASISCRSSQNLVYSDGN--TYLSWF 174

Db 121 SSGTGGSGGTGGTGGTDTIQLTQSPDSLAVSLGERATINCKSSQSVLYSSINKNYLAW 180

Qy 175 QORPGQSPRLIYKVSNRDGVPRDFSGSGGTDTFTLKISRVEAEDIGVYICMQGTHWPP 234

Db 181 QOKPGQPPKLLIYWASTRESGVPRDFSGSGGTDTFTLTISLQAEADVAVYICQ--YYSTP 239

Qy 235 RTFGQGTKEIKR 247

Db 240 YTFGQGTKEIKR 252

RESULT 14

US-11-337-300-276

; Sequence 276, Application US/11337300

; Publication No. US20060121580A1

; GENERAL INFORMATION:

; APPLICANT: Crucell Holland B.V.

; APPLICANT: De Kruif, Cornelis A.

; APPLICANT: van den Brink, Edward N.

; APPLICANT: Goudsmit, Jaap

; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof

; FILE REFERENCE: 0091 WO 00 ORD

; CURRENT APPLICATION NUMBER: US/11/337,300

; CURRENT FILING DATE: 2006-01-20

; NUMBER OF SEQ ID NOS: 478

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 276

; LENGTH: 254

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: SC03-052

US-11-337-300-276

Query Match

Best Local Similarity 55.9%; Score 782; DB 7; Length 254;

Matches 159; Conservative 33; Mismatches 53; Indels 8; Gaps 6;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGITHVRQAPGQGLEWMGWINNEN--T 58

Db 2 MAQVQLVQSGAEVKKPGESLKISKCKSGYGYFTYWGVRQMPGKGLWVGIIYPGDSSET 61

Qy 59 HYAQKFGQRTVMTTDTSTSTAYMELSLRSDDTAVYICARDPRGDDP--YWGQGTLLTV 116

Db 62 RYSPFQGVVTSADKSIINTAYLQWSSLKASDTAIYCA--GSGISTPMDVWGQGTLLTV 120

Qy 117 SSGGGSGGGSGGGGGS--EIVLTQSPSLPVTLPQASISCRSSQNLVYSDGN--TYLSWF 174

Db 121 SSGTGGSGGTGGTGGTDTIQLTQSPDSLAVSLGERATINCKSSQSVLYSSINKNYLAW 180

Qy 175 QORPGQSPRLIYKVSNRDGVPRDFSGSGGTDTFTLKISRVEAEDIGVYICMQGTHWPP 234

Db 181 QOKPGQPPKLLIYWASTRESGVPRDFSGSGGTDTFTLTISLQAEADVAVYICQ--YYSTP 239

Qy 235 RTFGQGTKEIKR 247

Db 240 YTFGQGTKEIKR 252

RESULT 15

US-11-337-300-274

; Sequence 274, Application US/11337300

; Publication No. US20060121580A1

; GENERAL INFORMATION:

; APPLICANT: Crucell Holland B.V.

; APPLICANT: ter Meulen, Jan H.

; APPLICANT: De Kruif, Cornelis A.

; APPLICANT: van den Brink, Edward N.

; APPLICANT: Goudsmit, Jaap

; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof

; FILE REFERENCE: 0091 WO 00 ORD

; CURRENT APPLICATION NUMBER: US/11/337,300

; CURRENT FILING DATE: 2006-01-20

; NUMBER OF SEQ ID NOS: 478

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 274

; LENGTH: 254

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: SC03-051

US-11-337-300-274

Query Match 55.8%; Score 780; DB 7; Length 254;

Best Local Similarity 62.5%; Pred. No. 2e-50;

Matches 158; Conservative 34; Mismatches 53; Indels 8; Gaps 6;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGITHVRQAPGQGLEWMGWINNEN--T 58

Db 2 MAQVQLVQSGAEVKKPGESLKISKCKSGYGYFTYWGVRQMPGKGLWVGIIYPGDSSET 61

Qy 59 HYAQKFGQRTVMTTDTSTSTAYMELSLRSDDTAVYICARDPRGDDP--YWGQGTLLTV 116

Db 62 RYSPFQGVVTSADKSIINTAYLQWSSLKASDTAIYCA--GSGISTPMDVWGQGTLLTV 120

Qy 117 SSGGGSGGGSGGGGGS--EIVLTQSPSLPVTLPQASISCRSSQNLVYSDGN--TYLSWF 174

Db 121 SSGTGGSGGTGGTGGTDTIQLTQSPDSLAVSLGERATINCKSSQSVLYSSINKNYLAW 180

Qy 175 QORPGQSPRLIYKVSNRDGVPRDFSGSGGTDTFTLKISRVEAEDIGVYICMQGTHWPP 234

Db 181 QOKPGQPPKLLIYWASTRESGVPRDFSGSGGTDTFTLTISLQAEADVAVYICQ--YYSTP 239

Qy 235 RTFGQGTKEIKR 247

Db 240 YTFGQGTKEIKR 252

Search completed: August 9, 2006, 13:30:53

Job time : 33 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:15:37 ; Search time 603 Seconds
(without alignments)
660.921 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVQLVQSGAEVKPGASV.....KVEIKRGLGLVYKDDDK 261

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 8366291 seqs, 1526956180 residues

Total number of hits satisfying chosen parameters: 8366291

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

1: /EMC_Celerra_SIDS3/ptodata/2/paa/US066 COMB.pap:*

2: /EMC_Celerra_SIDS3/ptodata/2/paa/US066 COMB.pap:*

3: /EMC_Celerra_SIDS3/ptodata/2/paa/US073 COMB.pap:*

4: /EMC_Celerra_SIDS3/ptodata/2/paa/US074 COMB.pap:*

5: /EMC_Celerra_SIDS3/ptodata/2/paa/US075 COMB.pap:*

6: /EMC_Celerra_SIDS3/ptodata/2/paa/US076 COMB.pap:*

7: /EMC_Celerra_SIDS3/ptodata/2/paa/US077 COMB.pap:*

8: /EMC_Celerra_SIDS3/ptodata/2/paa/US078 COMB.pap:*

9: /EMC_Celerra_SIDS3/ptodata/2/paa/US079 COMB.pap:*

10: /EMC_Celerra_SIDS3/ptodata/2/paa/US080 COMB.pap:*

11: /EMC_Celerra_SIDS3/ptodata/2/paa/US081 COMB.pap:*

12: /EMC_Celerra_SIDS3/ptodata/2/paa/US082 COMB.pap:*

13: /EMC_Celerra_SIDS3/ptodata/2/paa/US083 COMB.pap:*

14: /EMC_Celerra_SIDS3/ptodata/2/paa/US084 COMB.pap:*

15: /EMC_Celerra_SIDS3/ptodata/2/paa/US085 COMB.pap:*

16: /EMC_Celerra_SIDS3/ptodata/2/paa/US086 COMB.pap:*

17: /EMC_Celerra_SIDS3/ptodata/2/paa/US087 COMB.pap:*

18: /EMC_Celerra_SIDS3/ptodata/2/paa/US088 COMB.pap:*

19: /EMC_Celerra_SIDS3/ptodata/2/paa/US089 COMB.pap:*

20: /EMC_Celerra_SIDS3/ptodata/2/paa/US090 COMB.pap:*

21: /EMC_Celerra_SIDS3/ptodata/2/paa/US091 COMB.pap:*

22: /EMC_Celerra_SIDS3/ptodata/2/paa/US092 COMB.pap:*

23: /EMC_Celerra_SIDS3/ptodata/2/paa/US093 COMB.pap:*

24: /EMC_Celerra_SIDS3/ptodata/2/paa/US094 COMB.pap:*

25: /EMC_Celerra_SIDS3/ptodata/2/paa/US095 COMB.pap:*

26: /EMC_Celerra_SIDS3/ptodata/2/paa/US096 COMB.pap:*

27: /EMC_Celerra_SIDS3/ptodata/2/paa/US097 COMB.pap:*

28: /EMC_Celerra_SIDS3/ptodata/2/paa/US098 COMB.pap:*

29: /EMC_Celerra_SIDS3/ptodata/2/paa/US099 COMB.pap:*

30: /EMC_Celerra_SIDS3/ptodata/2/paa/US100 COMB.pap:*

31: /EMC_Celerra_SIDS3/ptodata/2/paa/US101 COMB.pap:*

32: /EMC_Celerra_SIDS3/ptodata/2/paa/US102 COMB.pap:*

33: /EMC_Celerra_SIDS3/ptodata/2/paa/US103 COMB.pap:*

34: /EMC_Celerra_SIDS3/ptodata/2/paa/US104 COMB.pap:*

35: /EMC_Celerra_SIDS3/ptodata/2/paa/US105 COMB.pap:*

36: /EMC_Celerra_SIDS3/ptodata/2/paa/US106 COMB.pap:*

37: /EMC_Celerra_SIDS3/ptodata/2/paa/US107 COMB.pap:*

38: /EMC_Celerra_SIDS3/ptodata/2/paa/US108 COMB.pap:*

39: /EMC_Celerra_SIDS3/ptodata/2/paa/US109 COMB.pap:*

40: /EMC_Celerra_SIDS3/ptodata/2/paa/US110 COMB.pap:*

41: /EMC_Celerra_SIDS3/ptodata/2/paa/US111 COMB.pap:*

42: /EMC_Celerra_SIDS3/ptodata/2/paa/US112 COMB.pap:*

43: /EMC_Celerra_SIDS3/ptodata/2/paa/US113 COMB.pap:*

44: /EMC_Celerra_SIDS3/ptodata/2/paa/US114 COMB.pap:*

45: /EMC_Celerra_SIDS3/ptodata/2/paa/US600 COMB.pap:*

46: /EMC_Celerra_SIDS3/ptodata/2/paa/US601 COMB.pap:*

47: /EMC_Celerra_SIDS3/ptodata/2/paa/US602 COMB.pap:*

48: /EMC_Celerra_SIDS3/ptodata/2/paa/US603 COMB.pap:*

49: /EMC_Celerra_SIDS3/ptodata/2/paa/US604 COMB.pap:*

50: /EMC_Celerra_SIDS3/ptodata/2/paa/US605 COMB.pap:*

51: /EMC_Celerra_SIDS3/ptodata/2/paa/US606 COMB.pap:*

52: /EMC_Celerra_SIDS3/ptodata/2/paa/US607 COMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1398	100.0	261	1	PCT-US05-04612-2 Sequence 2, Appli
2	1398	100.0	261	40	US-11-056-825-2 Sequence 2, Appli
3	1382	98.9	259	1	PCT-US05-04612-7 Sequence 7, Appli
4	1382	98.9	259	40	US-11-056-825-7 Sequence 7, Appli
5	1135	81.2	253	1	PCT-US01-19110-1964 Sequence 1964, Ap
6	1135	81.2	253	1	PCT-US02-36498-1964 Sequence 1964, Ap
7	1135	81.2	253	28	US-09-880-748-1964 Sequence 1964, Ap
8	1135	81.2	253	32	US-10-293-418-1964 Sequence 1964, Ap
9	1135	81.2	253	40	US-11-054-515-1964 Sequence 1964, Ap
10	1135	81.2	253	42	US-11-266-444-1964 Sequence 1964, Ap
11	1135	81.2	253	52	US-60-725-626-1964 Sequence 1964, Ap
12	1135	81.2	253	52	US-60-735-988-1964 Sequence 1964, Ap
13	1135	81.2	253	52	US-60-776-665-1964 Sequence 1964, Ap
14	1011.5	72.4	242	22	US-09-202-000-14 Sequence 14, Appl
15	988.5	70.7	252	37	US-10-762-593-589 Sequence 589, App
16	988.5	70.7	252	38	US-10-854-439-589 Sequence 589, App
17	983.5	70.4	286	35	US-10-535-764-184 Sequence 184, Appl
18	982.5	70.3	251	1	PCT-US00-19843-11 Sequence 11, Appl
19	982.5	70.3	251	30	US-10-331-783-11 Sequence 11, Appl
20	978.5	70.0	247	25	US-09-538-038-2343 Sequence 2343, Ap
21	978.5	70.0	247	25	US-09-538-038A-2343 Sequence 2343, Ap
22	963	68.9	249	1	PCT-US05-04612-4 Sequence 4, Appli
23	963	68.9	249	40	US-11-056-825-4 Sequence 4, Appli
24	952.5	68.1	524	35	US-10-551-504-293 Sequence 293, App
25	952	68.1	247	1	PCT-US05-04612-8 Sequence 8, Appli
26	952	68.1	247	40	US-11-056-825-8 Sequence 8, Appli
27	949.5	67.9	524	35	US-10-551-504-287 Sequence 287, App
28	947.5	67.8	524	35	US-10-551-504-2 Sequence 2, Appli
29	943	67.5	256	43	US-11-337-300-278 Sequence 278, App
30	939	67.2	281	35	US-10-535-764-186 Sequence 186, App
31	937.5	67.1	524	35	US-10-551-504-260 Sequence 260, App
32	933.5	66.8	274	32	US-10-257-864A-92 Sequence 92, Appl
33	933.5	66.8	274	36	US-10-645-085A-92 Sequence 92, Appl
34	933.5	66.8	524	35	US-10-551-504-254 Sequence 254, App
35	929.5	66.5	274	25	US-09-523-095A-26 Sequence 26, Appl
36	929.5	66.5	274	31	US-10-138-505-26 Sequence 26, Appl
37	929.5	66.5	274	32	US-10-257-864A-90 Sequence 90, Appl
38	929.5	66.5	274	33	US-10-399-518-119 Sequence 119, App
39	929.5	66.5	274	33	US-10-399-585-118 Sequence 118, App
40	929.5	66.5	274	36	US-10-645-085A-90 Sequence 90, Appl
41	928.5	66.4	274	25	US-09-523-095A-32 Sequence 32, Appl
42	928.5	66.4	274	31	US-10-138-505-32 Sequence 32, Appl
43	928.5	66.4	274	32	US-10-221-131-97 Sequence 97, Appl
44	928.5	66.4	274	33	US-10-399-518-121 Sequence 121, Appl
45	928.5	66.4	274	33	US-10-399-585-120 Sequence 120, App

ALIGNMENTS

RESULT 1
PCT-US05-04612-2
; Sequence 2, Application PC/TUS0504612
; GENERAL INFORMATION:

APPLICANT: Felding-Habermann, Brunhilde
APPLICANT: Janda, Kim D.
APPLICANT: Saven, Alan
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
FILE REFERENCE: SCRP-0043
CURRENT APPLICATION NUMBER: PCT/US05/04612
CURRENT FILING DATE: 2005-02-14
PRIOR FILING DATE: 2005-02-11
PRIOR FILING DATE: 2004-11-10
PRIOR FILING DATE: 2004-11-10
PRIOR FILING DATE: 2004-02-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent in version 3.3
SEQ ID NO 2
LENGTH: 261
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic Construct
PCT-US05-04612-2

Query Match 100.0%; Score 1398; DB 1; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-107;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
DB 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
QY 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
DB 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
QY 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
DB 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
QY 181 SPRRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
DB 181 SPRRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
QY 241 TKVEIKRGLGLVDYKDDDDK 261
DB 241 TKVEIKRGLGLVDYKDDDDK 261

RESULT 2
US-11-056-825-2
Sequence 2, Application US/11056825
GENERAL INFORMATION:
APPLICANT: Felding-Habermann, Brunhilde
APPLICANT: Janda, Kim D.
APPLICANT: Saven, Alan
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
FILE REFERENCE: SCRP-0042
CURRENT APPLICATION NUMBER: US/11/056,825
CURRENT FILING DATE: 2005-02-11
PRIOR FILING DATE: 2004-11-10
PRIOR FILING DATE: 2004-11-10
PRIOR FILING DATE: 2004-02-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent in version 3.3
SEQ ID NO 2
LENGTH: 261
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic Construct
US-11-056-825-2

Query Match 100.0%; Score 1398; DB 1; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-107;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
DB 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
QY 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
DB 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
QY 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
DB 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
QY 181 SPRRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
DB 181 SPRRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
QY 241 TKVEIKRGLGLVDYKDDDDK 261
DB 241 TKVEIKRGLGLVDYKDDDDK 261

Query Match 100.0%; Score 1398; DB 40; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-107;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
DB 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
QY 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
DB 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
QY 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
DB 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
QY 181 SPRRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
DB 181 SPRRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
QY 241 TKVEIKRGLGLVDYKDDDDK 261
DB 241 TKVEIKRGLGLVDYKDDDDK 261

RESULT 3
PCT-US05-04612-7
Sequence 7, Application PC/TUS0504612
GENERAL INFORMATION:
APPLICANT: Felding-Habermann, Brunhilde
APPLICANT: Janda, Kim D.
FILE REFERENCE: SCRP-0043
CURRENT APPLICATION NUMBER: PCT/US05/04612
CURRENT FILING DATE: 2005-02-14
PRIOR APPLICATION NUMBER: US 11/056,825
PRIOR FILING DATE: 2005-02-11
PRIOR APPLICATION NUMBER: US 60/626,726
PRIOR FILING DATE: 2004-11-10
PRIOR APPLICATION NUMBER: US 60/544,807
PRIOR FILING DATE: 2004-02-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent in version 3.3
SEQ ID NO 7
LENGTH: 259
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic Construct
PCT-US05-04612-7

Query Match 98.9%; Score 1382; DB 1; Length 259;
Best Local Similarity 99.6%; Pred. No. 5.6e-106;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
DB 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
QY 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
DB 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
QY 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
DB 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
QY 181 SPRRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
DB 181 SPRRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
QY 241 TKVEIKRGLGLVDYKDDDD 259

```
Db      241 TKVEIKRGLGLVDYKDD 259
|||||
RESULT 4
US-11-056-825-7
; Sequence 7, Application US/11056825
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Saven, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SSCP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR FILING DATE: 2004-11-10
; PRIOR FILING DATE: 2004-11-10
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO: 7
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-7
Query Match      98.9%; Score 1382; DB 40; Length 259;
Best Local Similarity 99.6%; Pred. No. 5.6e-106;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSNYGITVWRQAPGQGLEWMGWINNNTY 60
Db      1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSNYGITVWRQAPGQGLEWMGWINNNTY 60
Qy      61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Db      61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Qy      121 GSGGGGGGGGGSEIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFOQRPQG 180
Db      121 GSGGGGGGGGGSEIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFOQRPQG 180
Qy      181 SPRLIYKVNRRDGVPRDFSGSGSDTFTLKISRVEAEDIGVYICMQGTHWPTFTGQ 240
Db      181 SPRLIYKVNRRDGVPRDFSGSGSDTFTLKISRVEAEDIGVYICMQGTHWPTFTGQ 240
Qy      241 TKVEIKRGLGLVDYKDD 259
Db      241 TKVEIKRGLGLVDYKDD 259

RESULT 5
PCT-US01-19110-1964
; Sequence 1964, Application PC/TUS0119110
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523PCT
; CURRENT APPLICATION NUMBER: PCT/US01/19110
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25

Db      241 TKVEIKRGLGLVDYKDD 259
|||||
Query Match      98.9%; Score 1382; DB 40; Length 259;
Best Local Similarity 99.6%; Pred. No. 5.6e-106;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSNYGITVWRQAPGQGLEWMGWINNNTY 60
Db      1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSNYGITVWRQAPGQGLEWMGWINNNTY 60
Qy      61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Db      61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Qy      121 GSGGGGGGGGGSEIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFOQRPQG 180
Db      121 GSGGGGGGGGGSEIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFOQRPQG 180
Qy      181 SPRLIYKVNRRDGVPRDFSGSGSDTFTLKISRVEAEDIGVYICMQGTHWPTFTGQ 240
Db      181 SPRLIYKVNRRDGVPRDFSGSGSDTFTLKISRVEAEDIGVYICMQGTHWPTFTGQ 240
Qy      241 TKVEIKRGLGLVDYKDD 259
Db      241 TKVEIKRGLGLVDYKDD 259

RESULT 6
PCT-US02-36496-1964
; Sequence 1964, Application PC/TUS0236496
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523PCT2
; CURRENT APPLICATION NUMBER: PCT/US02/36496
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-36496-1964
Query Match      81.2%; Score 1135; DB 1; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy      3 QVQLVQSGAEVKPKGASVKSCASGYTFSNYGITVWRQAPGQGLEWMGWIN--NGNTHY 60
Db      1 QVQLQQSGAEVKPEPGASVKSCASGYTFTSYGISWVRQAPGQGLEWMGWSAYNGNTNY 60
Qy      61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTV 115
Db      61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTV 120
Qy      116 VSSGGGGGGGGGG--EIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSW 173
Db      121 VSSGGGGGGGGGG--EIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSW 180
Qy      174 FQRFQSPRLIYKVNRRDGVPRDFSGSGSDTFTLKISRVEAEDIGVYICMQGTHW 233
Db      181 FQRFQSPRLIYKVNRRDGVPRDFSGSGSDTFTLKISRVEAEDIGVYICMQGTHW 239
Qy      234 PTFGQGTKEIKR 247
Db      240 PTFGQGTKEIKR 253

RESULT 6
PCT-US02-36496-1964
; Sequence 1964, Application PC/TUS0236496
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523PCT2
; CURRENT APPLICATION NUMBER: PCT/US02/36496
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-36496-1964
Query Match      81.2%; Score 1135; DB 1; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy      3 QVQLVQSGAEVKPKGASVKSCASGYTFSNYGITVWRQAPGQGLEWMGWIN--NGNTHY 60
Db      1 QVQLQQSGAEVKPEPGASVKSCASGYTFTSYGISWVRQAPGQGLEWMGWSAYNGNTNY 60
Qy      61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTV 115
Db      61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTV 120
Qy      116 VSSGGGGGGGGGG--EIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSW 173
Db      121 VSSGGGGGGGGGG--EIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSW 180
Qy      174 FQRFQSPRLIYKVNRRDGVPRDFSGSGSDTFTLKISRVEAEDIGVYICMQGTHW 233
Db      181 FQRFQSPRLIYKVNRRDGVPRDFSGSGSDTFTLKISRVEAEDIGVYICMQGTHW 239
```

QY 234 PRTEGQGTKEIKR 247
| | | | | : | | | | |
Db 240 PFTFGQGTKEIKR 253

RESULT 7

US-09-880-748-1964
; Sequence 1964, Application US/09880748
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1964

Query Match 81.2%; Score 1135; DB 28; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

QY 3 QVQLVQSGAEVKPKGASVKVCKASGYTFSTNYGIVTVRQAPGQGLEWMGWIN--NGNTHY 60
| | | | | : | | | | | : | | | | | : | | | | |
Db 1 QVQLQSGAEVKPEKPCASVKVCKASGYTFSTYISWVRQAPGQGLEWMGWISAYNGNTNY 60
| | | | | : | | | | | : | | | | | : | | | | |
QY 61 AOKFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARNPYDSSGFFDYWGQGTMTV 115
| | | | | : | | | | | : | | | | | : | | | | |
Db 61 AOKLQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARNPYDSSGFFDYWGQGTMTV 120
| | | | | : | | | | | : | | | | | : | | | | |
QY 116 VSSGGSGGGGGGGGG--EIVLTQSPSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 173
| | | | | : | | | | | : | | | | | : | | | | |
Db 121 VSSGGSGGGGGGGGG--EIVLTQSPSLPVTLGQPASISCRSSQSLVYSDGNTYLNW 180
| | | | | : | | | | | : | | | | | : | | | | |
QY 174 FQRPQSPRLIYKVNRSRSDGVPDRFSGSGGTDTFTLKISRVEAEDIGVYYCMQGTHTP 233
| | | | | : | | | | | : | | | | | : | | | | |
Db 181 FQRPQSPRLIYKVNRSRSDGVPDRFSGSGGTDTFTLKISRVEAEDIGVYYCMQGTHTP 239
| | | | | : | | | | | : | | | | | : | | | | |
QY 234 PRTEGQGTKEIKR 247
| | | | | : | | | | |
Db 240 PFTFGQGTKEIKR 253

RESULT 8

US-10-293-418-1964
; Sequence 1964, Application US/10293418
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1964

Query Match 81.2%; Score 1135; DB 32; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

QY 3 QVQLVQSGAEVKPKGASVKVCKASGYTFSTNYGIVTVRQAPGQGLEWMGWIN--NGNTHY 60
| | | | | : | | | | | : | | | | | : | | | | |
Db 1 QVQLQSGAEVKPEKPCASVKVCKASGYTFSTYISWVRQAPGQGLEWMGWISAYNGNTNY 60
| | | | | : | | | | | : | | | | | : | | | | |
QY 61 AOKFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDD-----PYWGQGTMTV 115
| | | | | : | | | | | : | | | | | : | | | | |
Db 61 AOKLQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARNPYDSSGFFDYWGQGTMTV 120
| | | | | : | | | | | : | | | | | : | | | | |
QY 116 VSSGGSGGGGGGGGG--EIVLTQSPSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 173
| | | | | : | | | | | : | | | | | : | | | | |
Db 121 VSSGGSGGGGGGGGG--EIVLTQSPSLPVTLGQPASISCRSSQSLVYSDGNTYLNW 180
| | | | | : | | | | | : | | | | | : | | | | |
QY 174 FQRPQSPRLIYKVNRSRSDGVPDRFSGSGGTDTFTLKISRVEAEDIGVYYCMQGTHTP 233
| | | | | : | | | | | : | | | | | : | | | | |
Db 181 FQRPQSPRLIYKVNRSRSDGVPDRFSGSGGTDTFTLKISRVEAEDIGVYYCMQGTHTP 239
| | | | | : | | | | | : | | | | | : | | | | |
QY 234 PRTEGQGTKEIKR 247
| | | | | : | | | | |
Db 240 PFTFGQGTKEIKR 253

RESULT 9

US-11-054-515-1964
; Sequence 1964, Application US/11054515
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P3
; CURRENT APPLICATION NUMBER: US/11/054,515
; CURRENT FILING DATE: 2005-02-10
; PRIOR APPLICATION NUMBER: 60/543,296
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/580,347
; PRIOR FILING DATE: 2004-06-18
; PRIOR APPLICATION NUMBER: 10/293,418
; PRIOR FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964

Qy	176	ORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAEDIGVYYCMQGTHWPPER	235
Db	181	ORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAEDIGVYYCMQGTHWPPER	240
Qy	236	TFGQGTKVEIKR	247
Db	241	TFGQGTKVEIKR	252

Search completed: August 9, 2006, 13:26:30
Job time : 605 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:16:37 ; Search time 34 Seconds
(without alignments)
511.336 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVLVQSGAEVKKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 294602 seqs, 66610880 residues

Total number of hits satisfying chosen parameters: 294602

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:*

- 1: /EMC_Celerra_SID33/prodata/2/paa/PCT_NEW_COMB.pap:*
- 2: /EMC_Celerra_SID33/prodata/2/paa/US06_NEW_COMB.pap:*
- 3: /EMC_Celerra_SID33/prodata/2/paa/US07_NEW_COMB.pap:*
- 4: /EMC_Celerra_SID33/prodata/2/paa/US08_NEW_COMB.pap:*
- 5: /EMC_Celerra_SID33/prodata/2/paa/US09_NEW_COMB.pap:*
- 6: /EMC_Celerra_SID33/prodata/2/paa/US10_NEW_COMB.pap:*
- 7: /EMC_Celerra_SID33/prodata/2/paa/US11_NEW_COMB.pap:*
- 8: /EMC_Celerra_SID33/prodata/2/paa/US60_NEW_COMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1135	81.2	253	8	US-60-797-351-1964 Sequence 1964, Ap
2	1135	81.2	253	8	US-60-814-869-1964 Sequence 1964, Ap
3	1135	81.2	253	8	US-60-815-559-1964 Sequence 1964, Ap
4	1135	81.2	253	8	US-60-834-152-1964 Sequence 1964, Ap
5	924.5	66.1	244	8	US-60-797-351-1991 Sequence 1991, Ap
6	924.5	66.1	244	8	US-60-814-869-1991 Sequence 1991, Ap
7	924.5	66.1	244	8	US-60-815-559-1991 Sequence 1991, Ap
8	924.5	66.1	244	8	US-60-834-152-1991 Sequence 1991, Ap
9	919.5	65.8	277	7	US-11-177-495-3 Sequence 3, Appli
10	906	64.8	251	8	US-60-797-351-1238 Sequence 1238, Ap
11	906	64.8	251	8	US-60-814-869-1238 Sequence 1238, Ap
12	906	64.8	251	8	US-60-815-559-1238 Sequence 1238, Ap
13	906	64.8	251	8	US-60-834-152-1238 Sequence 1238, Ap
14	899	64.3	251	8	US-60-797-351-1921 Sequence 1921, Ap
15	899	64.3	251	8	US-60-814-869-1921 Sequence 1921, Ap
16	899	64.3	251	8	US-60-815-559-1921 Sequence 1921, Ap
17	899	64.3	251	8	US-60-834-152-1921 Sequence 1921, Ap
18	898.5	64.3	276	6	US-10-536-834A-95 Sequence 95, Appl
19	893.5	63.9	250	8	US-60-797-351-62 Sequence 62, Appl
20	893.5	63.9	250	8	US-60-814-869-62 Sequence 62, Appl
21	893.5	63.9	250	8	US-60-815-559-62 Sequence 62, Appl
22	893.5	63.9	250	8	US-60-834-152-62 Sequence 62, Appl
23	893	63.9	251	8	US-60-797-351-85 Sequence 85, Appl
24	893	63.9	251	8	US-60-814-869-85 Sequence 85, Appl
25	893	63.9	251	8	US-60-815-559-85 Sequence 85, Appl

26	893	63.9	251	8	US-60-834-152-85 Sequence 85, Appl
27	891.5	63.8	248	8	US-60-797-351-1446 Sequence 1446, Ap
28	891.5	63.8	248	8	US-60-814-869-1446 Sequence 1446, Ap
29	891.5	63.8	248	8	US-60-815-559-1446 Sequence 1446, Ap
30	891.5	63.8	248	8	US-60-834-152-1446 Sequence 1446, Ap
31	890.5	63.7	260	8	US-60-797-351-1039 Sequence 1039, Ap
32	890.5	63.7	260	8	US-60-814-869-1039 Sequence 1039, Ap
33	890.5	63.7	260	8	US-60-815-559-1039 Sequence 1039, Ap
34	890.5	63.7	260	8	US-60-834-152-1039 Sequence 1039, Ap
35	887	63.4	251	8	US-60-797-351-264 Sequence 264, App
36	887	63.4	251	8	US-60-814-869-264 Sequence 264, App
37	887	63.4	251	8	US-60-815-559-264 Sequence 264, App
38	887	63.4	251	8	US-60-834-152-264 Sequence 264, App
39	886	63.4	251	8	US-60-797-351-178 Sequence 178, App
40	886	63.4	251	8	US-60-814-869-178 Sequence 178, App
41	886	63.4	251	8	US-60-815-559-178 Sequence 178, App
42	886	63.4	251	8	US-60-834-152-178 Sequence 178, App
43	885	63.3	251	8	US-60-797-351-234 Sequence 234, App
44	885	63.3	251	8	US-60-814-869-234 Sequence 234, App
45	885	63.3	251	8	US-60-815-559-234 Sequence 234, App

ALIGNMENTS

RESULT 1
US-60-797-351-1964
; Sequence 1964, Application US/60797351
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind to B Lymphocyte Stimulator;
; FILE REFERENCE: PF523PP11
; CURRENT APPLICATION NUMBER: US/60797,351
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-797-351-1964

Query Match 81.2%; Score 1135; DB 8; Length 253;
Best Local Similarity 86.2%; Pred. No. 9.2e-81;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy	3	QVQLVQSGAEVKKPGASVKVCSGYTFTSNGITWVROAPGQGLEWMGWIN--NGNTHY	60
Db	1	QVQLVQSGAEVKKPGASVKVCSGYTFTSNGITWVROAPGQGLEWMGWIN--NGNTHY	60
Qy	61	AQKFGQRTVTMTDTSTSTAYMELRSLRSDDTAVYVCARDPRGDDE-----PYWGQGLT	115
Db	61	AQKFGQRTVTMTDTSTSTAYMELRSLRSDDTAVYVCARDPRGDDE-----PYWGQGLT	115
Qy	116	VSSGGGGGGGGGGGGGG--EIVLTQSLPLSLPVTIQLQGPASISCRSSQNLVYSDGNTYLSW	173
Db	121	VSSGGGGGGGGGGGGGG--EIVLTQSLPLSLPVTIQLQGPASISCRSSQNLVYSDGNTYLSW	180
Qy	174	FOQRPGQSPRLIYKVNRDSCVPPFSGSGSGTFTLKISRVEAEDIGVYCMQGTHTP	233
Db	181	FOQRPGQSPRLIYKVNRDSCVPPFSGSGSGTFTLKISRVEAEDIGVYCMQGTHTP	239
Qy	234	PRTFGGGTGKVEIKR	247
Db	240	PRTFGGGTGKVEIKR	253

RESULT 2
US-60-814-869-1964
; Sequence 1964, Application US/60814869
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind to B Lymphocyte Stimulator;
; FILE REFERENCE: PF523PP12

; CURRENT APPLICATION NUMBER: US/60/814,869
; CURRENT FILING DATE: 2006-06-20
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-814-869-1964

Query Match 81.2%; Score 1135; DB 8; Length 253;
Best Local Similarity 86.2%; Pred. No. 9.2e-81;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;
QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGITWVRQAPQGQLEWMGWIN--NGNTHY 60
DB 1 QVQLQSGAEVKKPGASVKVSCKASGYTFSTNYGISWVRQAPQGQLEWMGWISAYNGNTNY 60
QY 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDD-----PYWGQGTILVT 115
DB 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDD-----PYWGQGTILVT 120
QY 116 VSSGGSGGGGGGGGGG--EIVLTQSPPLSLPVTLGQSPASISCRSSQNLVYSDGNTYLSW 173
DB 121 VSSGGSGGGGGGGGGG--EIVLTQSPPLSLPVTLGQSPASISCRSSQNLVYSDGNTYLSW 180
QY 174 FQRPQSPRLIYKVNRSQVDPDRFSGSGGTDFTLKISRVEADIGVYYCMQGTWHP 233
DB 181 FQRPQSPRLIYKVNRSQVDPDRFSGSGGTDFTLKISRVEADIGVYYCMQGTWHP 239
QY 234 PRTFGQGTKEIKR 247
DB 240 PFTFGQGTKEIKR 253

RESULT 3
US-60-815-559-1964
; Sequence 1964, Application US/60815559
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP13
; CURRENT APPLICATION NUMBER: US/60/815,559
; CURRENT FILING DATE: 2006-06-22
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-815-559-1964

Query Match 81.2%; Score 1135; DB 8; Length 253;
Best Local Similarity 86.2%; Pred. No. 9.2e-81;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;
QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGITWVRQAPQGQLEWMGWIN--NGNTHY 60
DB 1 QVQLQSGAEVKKPGASVKVSCKASGYTFSTNYGISWVRQAPQGQLEWMGWISAYNGNTNY 60
QY 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDD-----PYWGQGTILVT 115
DB 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDD-----PYWGQGTILVT 120
QY 116 VSSGGSGGGGGGGGGG--EIVLTQSPPLSLPVTLGQSPASISCRSSQNLVYSDGNTYLSW 173
DB 121 VSSGGSGGGGGGGGGG--EIVLTQSPPLSLPVTLGQSPASISCRSSQNLVYSDGNTYLSW 180
QY 174 FQRPQSPRLIYKVNRSQVDPDRFSGSGGTDFTLKISRVEADIGVYYCMQGTWHP 233
DB 181 FQRPQSPRLIYKVNRSQVDPDRFSGSGGTDFTLKISRVEADIGVYYCMQGTWHP 239
QY 234 PRTFGQGTKEIKR 247
DB 240 PFTFGQGTKEIKR 253

RESULT 4
US-60-834-152-1964
; Sequence 1964, Application US/60834152
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP14
; CURRENT APPLICATION NUMBER: US/60/834,152
; CURRENT FILING DATE: 2006-07-31
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-834-152-1964

Query Match 81.2%; Score 1135; DB 8; Length 253;
Best Local Similarity 86.2%; Pred. No. 9.2e-81;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;
QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGITWVRQAPQGQLEWMGWIN--NGNTHY 60
DB 1 QVQLQSGAEVKKPGASVKVSCKASGYTFSTNYGISWVRQAPQGQLEWMGWISAYNGNTNY 60
QY 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDD-----PYWGQGTILVT 115
DB 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDD-----PYWGQGTILVT 120
QY 116 VSSGGSGGGGGGGGGG--EIVLTQSPPLSLPVTLGQSPASISCRSSQNLVYSDGNTYLSW 173
DB 121 VSSGGSGGGGGGGGGG--EIVLTQSPPLSLPVTLGQSPASISCRSSQNLVYSDGNTYLSW 180
QY 174 FQRPQSPRLIYKVNRSQVDPDRFSGSGGTDFTLKISRVEADIGVYYCMQGTWHP 233
DB 181 FQRPQSPRLIYKVNRSQVDPDRFSGSGGTDFTLKISRVEADIGVYYCMQGTWHP 239
QY 234 PRTFGQGTKEIKR 247
DB 240 PFTFGQGTKEIKR 253

RESULT 5
US-60-797-351-1991
; Sequence 1991, Application US/60797351
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP11
; CURRENT APPLICATION NUMBER: US/60/797,351
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1991
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-797-351-1991

Query Match 66.1%; Score 924.5; DB 8; Length 244;
Best Local Similarity 71.1%; Pred. No. 1.8e-64;
Matches 177; Conservative 34; Mismatches 29; Indels 9; Gaps 4;
QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGITWVRQAPQGQLEWMGWIN--NGNTHY 60
DB 1 EYQLVQSGGGLVPGGSLRLSCAASGFTFSSYMMNWVRQAPGKGLVWSSISSSSHHY 60
QY 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPYWGQGTILVTSVSSG 120
DB 61 ADSVKGRFTISRDNKNSLYLQWNSLRADTAIVYICARVHSSGS---WGQGTILVTSVSSG 117
QY 121 GSGGGSGGGGGG--EIVLTQSPPLSLPVTLGQSPASISCRSSQNLVYSDGNTYLSWFOORP 178

Db 118 GSGGGGGGGGSGGALDVVMTQSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 177
QY 179 GQSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 238
Db 178 GQSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 235
QY 239 QGTVKVEIKR 247
Db 236 QGTVKVEIKR 244

RESULT 6
US-60-814-869-1991
; Sequence 1991, Application US/60814869
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP12
; CURRENT APPLICATION NUMBER: US/60/814,869
; NUMBER OF SEQ ID NOS: 2006-06-20
; SEQ ID NO 1991
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-814-869-1991

Query Match 66.1%; Score 924.5; DB 8; Length 244;
Best Local Similarity 71.1%; Pred. No. 1.8e-64;
Matches 177; Conservative 34; Mismatches 29; Indels 9; Gaps 4;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSSNYGIVTVRQAPGQGLEWMGWINNNGTH--Y 60
Db 1 EVQLVQSGGGLVKPGGSLRLSCAASGFTFSYSMNWVRQAPGKGLWVSSISSSSHIYY 60
QY 61 AQKFGQRTVMTDTSTAYMELRSLRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGG 120
Db 61 ADSVKGRFTISRDNKNSLYLQMSLRRAEDTAVYICARVHSSGS---WGQGLTVTVSSGG 117
QY 121 GSGGGGGGGGSGG--EIVLTQSPSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 178
Db 118 GSGGGGGGGGSGGALDVVMTQSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 177
QY 179 GQSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 238
Db 178 GQSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 235
QY 239 QGTVKVEIKR 247
Db 236 QGTVKVEIKR 244

RESULT 7
US-60-815-559-1991
; Sequence 1991, Application US/60815559
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP13
; CURRENT APPLICATION NUMBER: US/60/815,559
; NUMBER OF SEQ ID NOS: 2006-06-22
; SEQ ID NO 1991
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-815-559-1991

Query Match 66.1%; Score 924.5; DB 8; Length 244;
Best Local Similarity 71.1%; Pred. No. 1.8e-64;
Matches 177; Conservative 34; Mismatches 29; Indels 9; Gaps 4;
QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSSNYGIVTVRQAPGQGLEWMGWINNNGTH--Y 60

Db 1 EVQLVQSGGGLVKPGGSLRLSCAASGFTFSYSMNWVRQAPGKGLWVSSISSSSHIYY 60
QY 61 AQKFGQRTVMTDTSTAYMELRSLRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGG 120
Db 61 ADSVKGRFTISRDNKNSLYLQMSLRRAEDTAVYICARVHSSGS---WGQGLTVTVSSGG 117
QY 121 GSGGGGGGGGSGG--EIVLTQSPSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 178
Db 118 GSGGGGGGGGSGGALDVVMTQSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 177
QY 179 GQSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 238
Db 178 GQSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 235
QY 239 QGTVKVEIKR 247
Db 236 QGTVKVEIKR 244

RESULT 8
US-60-834-152-1991
; Sequence 1991, Application US/60834152
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP14
; CURRENT APPLICATION NUMBER: US/60/834,152
; CURRENT FILING DATE: 2006-07-31
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1991
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-834-152-1991

Query Match 66.1%; Score 924.5; DB 8; Length 244;
Best Local Similarity 71.1%; Pred. No. 1.8e-64;
Matches 177; Conservative 34; Mismatches 29; Indels 9; Gaps 4;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSSNYGIVTVRQAPGQGLEWMGWINNNGTH--Y 60
Db 1 EVQLVQSGGGLVKPGGSLRLSCAASGFTFSYSMNWVRQAPGKGLWVSSISSSSHIYY 60
QY 61 AQKFGQRTVMTDTSTAYMELRSLRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGG 120
Db 61 ADSVKGRFTISRDNKNSLYLQMSLRRAEDTAVYICARVHSSGS---WGQGLTVTVSSGG 117
QY 121 GSGGGGGGGGSGG--EIVLTQSPSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 178
Db 118 GSGGGGGGGGSGGALDVVMTQSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 177
QY 179 GQSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 238
Db 178 GQSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 235
QY 239 QGTVKVEIKR 247
Db 236 QGTVKVEIKR 244

RESULT 9
US-11-177-495-3
; Sequence 3, Application US/11177495
; GENERAL INFORMATION:
; APPLICANT: RAPAPORT FAMILY INSTITUTE FOR RESEARCH
; TITLE OF INVENTION: METHODS OF DETECTING HAPTOGLOBIN PHENOTYPE
; FILE REFERENCE: P-7018-US
; CURRENT APPLICATION NUMBER: US/11/177,495
; CURRENT FILING DATE: 2005-07-11
; PRIOR APPLICATION NUMBER: USP 60/586,733
; PRIOR FILING DATE: 2004-07-12
; NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn version 3.3

SEQ ID NO 3

LENGTH: 277

TYPE: PRN

ORGANISM: Human

US-11-177-495-3

Query Match

Best Local Similarity 65.8%; Score 919.5; DB 7; Length 277;

Matches 175; Conservative 39; Mismatches 29; Indels 9; Gaps 4;

QY 1 MAQVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGQGLEWMGW--NGNT 58

DB 1 MAEVQSQSGTELVKPGASVKVSCKASGYTFSTNYGIVTWRQAPGQGLEWMGW--NGNT 60

QY 59 HYAQFGQGRVTMTTDTSTAYMELSLRSDTAVYICARDPRGDDE---PYWGQGLVLT 115

DB 61 KFNFRNKATLTVDKSNYATMQVSSLTSDSAVYICARD---DDSAWYAWGQGLVLT 117

QY 116 VSSGGGSGGGGGGGGSEIVLTQSPISLPLVTLGQPASISCRSSQNLVYSDGNTYLSWFQ 175

DB 118 VSSGGGSGGGGGGGGSDVVMTQTPLSLPVSGLDQASISCRSSQSIHVSNGNTYLEWYL 177

QY 176 QPCGSPRLIYKVNRSQGVDPDRFSGSGGTDFTLKISRVEADIGVYCMQGTHTWPPR 235

DB 178 QKPGSPKLLIYKVNRFSGVDPDRFSGSGGTDFTLKISRVEADIGVYCFQGSHPVW 236

QY 236 TFGGQTKVEIKR 247

DB 237 TFGGQTKLEIKR 248

RESULT 10

US-60-797-351-1238

Sequence 1238, Application US/60797351

GENERAL INFORMATION:

APPLICANT: Ruben et al.

TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator

FILE REFERENCE: PF523PP11

CURRENT APPLICATION NUMBER: US/60/797,351

CURRENT FILING DATE: 2006-05-04

NUMBER OF SEQ ID NOS: 3247

SEQ ID NO 1238

LENGTH: 251

TYPE: PRN

ORGANISM: Homo sapiens

US-60-797-351-1238

Query Match

Best Local Similarity 64.8%; Score 906; DB 8; Length 251;

Matches 178; Conservative 29; Mismatches 26; Indels 30; Gaps 5;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGQGLEWMGW--INNNGNTHY 60

DB 1 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGQGLEWMGW--INNNGNTHY 60

QY 61 AQKFGQGRVTMTTDTSTAYMELSLRSDTAVYICARDPRGDDEPY-----PYDPLTAYTFQYFG 107

DB 61 AQKFGQGRVTMTTDTSTAYMELSLRSDTAVYICARDPRGDDEPY-----PYDPLTAYTFQYFG 113

QY 108 -WGQGLTVTVSSGGGGGGGGGGGSEIVLTQSPISLPLVTLGQPASISCRSSQNLVY 164

DB 114 NWGQGLTVTVSSGGGGGGGGGGGSEIVLTQSPISLPLVTLGQPASISCRSSQNLVY 171

QY 165 SDGNTYLSWFOQPGQSPRLIYKVNRSQGVDPDRFSGSGGTDFTLKISRVEADIGVY 224

DB 172 ---SSNLAWFOQPGQAPRLIYGASTRFTGTPARFSGSGGTDFSLTISLSQSDPAVY 228

QY 225 YCMQGTHTWPPRTFGQTKVEIKR 247

DB 229 YCQYNNWPPITFGQTKLEIKR 251

RESULT 11

US-60-814-869-1238

Sequence 1238, Application US/60814869

GENERAL INFORMATION:

APPLICANT: Ruben et al.

TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator

FILE REFERENCE: PF523PP12

CURRENT APPLICATION NUMBER: US/60/814,869

CURRENT FILING DATE: 2006-06-20

NUMBER OF SEQ ID NOS: 3247

SEQ ID NO 1238

LENGTH: 251

TYPE: PRN

ORGANISM: Homo sapiens

US-60-814-869-1238

Query Match

Best Local Similarity 64.8%; Score 906; DB 8; Length 251;

Matches 178; Conservative 29; Mismatches 26; Indels 30; Gaps 5;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGQGLEWMGW--INNNGNTHY 60

DB 1 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGQGLEWMGW--INNNGNTHY 60

QY 61 AQKFGQGRVTMTTDTSTAYMELSLRSDTAVYICARDPRGDDEPY-----PYDPLTAYTFQYFG 107

DB 61 AQKFGQGRVTMTTDTSTAYMELSLRSDTAVYICARDPRGDDEPY-----PYDPLTAYTFQYFG 113

QY 108 -WGQGLTVTVSSGGGGGGGGGGGSEIVLTQSPISLPLVTLGQPASISCRSSQNLVY 164

DB 114 NWGQGLTVTVSSGGGGGGGGGGGSEIVLTQSPISLPLVTLGQPASISCRSSQNLVY 171

QY 165 SDGNTYLSWFOQPGQSPRLIYKVNRSQGVDPDRFSGSGGTDFTLKISRVEADIGVY 224

DB 172 ---SSNLAWFOQPGQAPRLIYGASTRFTGTPARFSGSGGTDFSLTISLSQSDPAVY 228

QY 225 YCMQGTHTWPPRTFGQTKVEIKR 247

DB 229 YCQYNNWPPITFGQTKLEIKR 251

RESULT 12

US-60-815-559-1238

Sequence 1238, Application US/60815559

GENERAL INFORMATION:

APPLICANT: Ruben et al.

TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator

FILE REFERENCE: PF523PP13

CURRENT APPLICATION NUMBER: US/60/815,559

CURRENT FILING DATE: 2006-06-22

NUMBER OF SEQ ID NOS: 3247

SEQ ID NO 1238

LENGTH: 251

TYPE: PRN

ORGANISM: Homo sapiens

US-60-815-559-1238

Query Match

Best Local Similarity 64.8%; Score 906; DB 8; Length 251;

Matches 178; Conservative 29; Mismatches 26; Indels 30; Gaps 5;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGQGLEWMGW--INNNGNTHY 60

DB 1 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGQGLEWMGW--INNNGNTHY 60

QY 61 AQKFGQGRVTMTTDTSTAYMELSLRSDTAVYICARDPRGDDEPY-----PYDPLTAYTFQYFG 107

DB 61 AQKFGQGRVTMTTDTSTAYMELSLRSDTAVYICARDPRGDDEPY-----PYDPLTAYTFQYFG 113

QY 108 -WGQGLTVTVSSGGGGGGGGGGGSEIVLTQSPISLPLVTLGQPASISCRSSQNLVY 164

DB 114 NWGQGLTVTVSSGGGGGGGGGGGSEIVLTQSPISLPLVTLGQPASISCRSSQNLVY 171

165	QY	SDGNTYLSWFOORCGSPRLLIYKVNSRDSGVPRPFCSGSGTDTFLIKSRVEARDIGVY	224
172	Db	---SSNLAWFOOKGQAPRLLIYGASTRTGIPARFCSGSGTDFSLTISLQSDFAVY	228
225	QY	YCMQGTWPPRTFGQGTKVEIKR	247
229	Db	YCOQYNWPPRTFGQGTLEIKR	251

RESULT 13
US-60-834-152-1238
; Sequence 1238, Application US/60834152
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulat
; FILE REFERENCE: PFS23PP14
; CURRENT APPLICATION NUMBER: US/60/834,152
; CURRENT FILING DATE: 2006-07-31
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1238
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-834-152-1238

Query Match 64.8%; Score 906; DB 8; Length 251;
Best Local Similarity 67.7%; Pred. No. 4.9e-63;
Matches 178; Conservative 29; Mismatches 26; Gaps 5;
Indels 30; Gaps 5;

	Qy	3	OVOLVQSGBAEVKKPGASVKYSCKASGYTFSSNYGHTWVRQAPOGLEWGM--INNGNTHY	60
	Dd	1	OVOLVQSGBAEVKKPGASVKYSCKASGYTFSHGISWRVQAPOGLEWGWNSAHDNTKY	60
	Qy	61	AQRFQGRVTMTTDTSTSTAYMELSLRSDDTAIVYYCARDPRGDDEPY-----	107
	Dd	61	AQRFQGRVTMTTDASTSTAYMELSLRSDDTAIVYYCAR-----PYDPLTAVTFQYFG	113
	Qy	108	-WGQGTLTVTVSSGGGGSGGGSGGGGGGS--EIVLTQSPLSLPTVIGQPASISCRSSQNLYV	164
	Dd	114	NWGQGTLTVTVSSGGGGSGGGSGGGGGSALETTLTQSPATLVSPFGERVTLSCRASQSV--	171
	Qy	165	SDGNITYLWFQQRPQGPQRLLIYKVSNRDSGVDPDFSGSGGTDFTLIKISRVEAEDIGVY	224
	Dd	172	---SSNLAWFQQRPQGPQRLLIYGASTRTFTGIPIAPFSGSGGTDFSLTISSLQSEDFAVY	228
	Qy	225	YCMQGHWPDPRTFGQTKVBIEKR	247
	Dd	229	YCQOYNWPPITFGQTRLEIKR	251

```

RESULT 14
US-60-797-351-1921
; Sequence 1921, Application US/60797351
; GENERAL INFORMATION:
; APPLICANT: Ruben et al
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulat
; FILE REFERENCE: PF523PP11
; CURRENT APPLICATION NUMBER: US/60/797,351
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1921
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-797-351-1921

```

Query Match	64.3%	Score 899;	DB 8;	Length 251;
Best Local Similarity	68.9%;	Pred. No. 1.7e-62;		
Matches 177;	Conservative	27;	Mismatches 35;	Indels 18;
			Gaps 5;	

Qy 3 QVQLVQSGAEVKKPGASVKYCKASGYTSPSYGIIITWRQAPQGLEWGWIN--NGNTHY 60

Qy	61	AQKFGQGRVTMTTDTSTAYMELRSRSDDTAVYYCAR	110	-----DPRGDDPEY-----WGQ
Db	61	AQKLGQGRVTMTTDTSTAYMELRSRSDDTAVYYCAR	120	-----SSSGYYGYGMDVWGR
Qy	111	GTLLVTYSSGGGGGGGGSEIVLTQSPLSLPVTLLGQAPASISCRSSONLVIYSDGNTY	170	
Db	121	GTITVTYSSGGGGGGGGGGSDIQMTQSPSFLSASVGDRVTITCRASQGI-----NNY	175	
Qy	171	LSWFQQRPGOSPRLLIYKVSNRDSDGVPDRFSGSGSGTDFTLTKISRVEADIGVYCMQGT	230	
Db	176	LAWVQKQKPGRAPKLLIYAASLSQGVSPRFGSGSGTDFTLTISSLPQDPFATYYCLQDS	235	
Qy	231	HWPPRTFGQGTAKVBEIKR	247	
Db	236	DY-PLTFGGGTGLEIKR	251	

RESULT 15

```

RESULTS 13
US-60-814-869-1921
; Sequence 1921, Application US/60814869
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP12
; CURRENT APPLICATION NUMBER: US/60/814,869
; CURRENT FILING DATE: 2006-06-20
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1921
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-814-869-1921

```

Query Match 64.3%: Score 899: DB 8: Length 251:

Query match 64.3%; score 899; DB % mengacu 231;
Best Local Similarity 68.9%; Pred. No. 1.7e-62;
Matches 177; Conservative 27; Mismatches 35; Indels 18; Gaps 5;

Qy 3 QVOLVQSGAEVKKPGASVKYCKASGYTFSNYGIITWRQAPGGGLEWMGWIN--NGNTHY 60

Db 1 QVOLVQSGAEVKKPGASVKYCKASGYTFSNYGIISWRQAPGGGLEWMGWISAYNGNTNY 60

	61	AQKPGRVMTTDTSTSTAYMELRSLRSDDTAVYYCAR	-----DPRGDDEPY	WGQ	110
Qy		:	:		
	61	AOKLGRVMTTDTSTSTAYMELRSLRSDDTAVYYCAR	TLSYSSSSGGYYYGMDVWGR		120
Dd		:	:		

Qy	111	GTLVTVSSGGGSGGGSGGGSEIVLTQSPPLSLPVTVLTQGPASISCRSSQNQLVYSDGNTY	170
Db	121	GTITVTYSSGGGSGGGGSGGGSDIOMTQSPSLFASGDRVITITCRASOGI	175

Qy	171	176	235	230
Qy	LSWFQQRPGQSPRRLLIKVSNRDSGVDPDRFGSGSGTPTLTKLSRVEAEDIGVYCMQGT			
Db		LWYQKQGRAPKLLIIYAASSLQSGVSPRFGSGSGTPTLTLSLQPEDFATYVCLQDS		

Qy	231	HWPRTFGQTKVEIKR	247
		:	
Db	236	DY-PLTEGGGKLEIKR	251

Search completed: August 9, 2006, 13:27:10
Job time : 35 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:09:51 ; Search time 40 Seconds
(without alignment)
627.814 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVQLVQSGAEVKKPGASV.....KVEIKRGLGLVDYKDDDDK 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_80.*

2: Pirl.*

3: Pirl3.*

4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	808	57.8	249	2 S41374	single chain Fv an
2	676.5	48.4	268	2 A56446	Ig heavy chain V r
3	660	47.2	233	2 J05322	p53 specific singl
4	565	40.4	114	2 B49002	Ig kappa chain V r
5	562	40.2	114	2 S49572	Ig kappa chain pre
6	561.5	40.2	133	2 S23230	Ig kappa chain pre
7	560.5	40.1	133	1 K2HURP	Ig kappa chain pre
8	553	39.6	140	2 S22658	Ig kappa chain pre
9	552.5	39.5	133	1 A24452	Ig kappa chain pre
10	551.5	39.4	133	2 S40324	Ig kappa chain V r
11	550.5	39.4	122	2 S40338	Ig kappa chain - h
12	547.5	39.2	142	2 S22902	Ig kappa chain V r
13	546.5	39.1	133	2 S42611	HUNVK protein prec
14	535.5	38.3	118	2 S40374	Ig kappa chain - h
15	529.5	37.9	132	2 S40322	Ig kappa chain - h
16	521.5	37.3	160	2 P0105	anti-PR2 erythrocy
17	521	37.3	120	2 S42268	Ig kappa chain V r
18	513	36.7	122	2 S42267	Ig kappa chain V r
19	509.5	36.4	122	2 S36271	Ig heavy chain V r
20	507	36.3	134	2 S40376	Ig kappa chain - h
21	505.5	36.2	129	2 S36260	Ig heavy chain V r
22	501.5	35.9	126	2 S40312	Ig kappa chain - h
23	496	35.5	124	2 S19665	Ig heavy chain V r
24	487	34.8	118	2 S36265	Ig heavy chain V r
25	486	34.8	135	2 S49530	anti-Sm antibody V
26	478.5	34.2	142	2 S42483	Ig heavy chain V r
27	473.5	33.9	125	2 S68170	Ig heavy chain V r
28	471.5	33.7	128	2 S40373	Ig kappa chain - h
29	471.5	33.7	219	2 S52028	Ig kappa chain - m

RESULT 1

S41374

single chain Fv antibody - mouse

C:Species: Mus musculus (house mouse)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995

C:Accession: S41374

R:Artsaenko, O.; Weiler, E.W.; Muentz, K.; Conrad, U.

submitted to the EMBL Data Library, January 1994

A:Description: Construction and functional characterization of a single chain Fv antibody

A:Reference number: S41374

A:Accession: S41374

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-249 <ART>

A:Cross-references: UNIPARC:UPI000017C74A; EMBL:Z29480

Query Match

Best Local Similarity 57.8%; Score 808; DB 2; Length 249;

Matches 161; Conservative 33; Mismatches 50; Indels 4; Gaps 3;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASQGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNTHY 60

DB 1 QVQLVQSGAEVKKPGASVKLSCTASGFNPKDDYIHVWQRPKEGLEWIAIPASGNVKY 60

QY 61 AOKFGQRTVMTDTSTAYMELRLSRDSTAVVYCA-RDPRGDDPEYWGQGLTVTVSSG 119

DB 61 VPRFQDKATITADTSNTAYILLSSLTSEDYAVVYCARRDYLTSLGYWGQGSTVTVSSR 120

QY 120 GGGSGGGSGGGGGSHVLTQSPFLSLPVTILGPASISCRSSQNLYSDGNNTYLSWFOQRP 179

DB 121 GGGSGGGSGGGGGSHVLTQSPFPVVPVPGESVSISSCRSSKLLYSDGSLFWFLQRP 180

QY 180 QSPRLIYKVNDRDGVDPFRFSGSGSGTDFTLKISRVEAEDGIVYCMQGTWHPPTFGQ 239

DB 181 QSPQLLIYRMSNLASGVDPFRFSGSGSGTFTLIRISREAEADVGVVYCMQHREY-PLTFGA 239

QY 240 GTKVEIKR 247

DB 240 GTKLEIKR 247

RESULT 2

A56446

Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/C)

C:Species: Mus musculus (house mouse)

C>Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 16-Aug-1996

C:Accession: A56446

R:Tang, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.

J. Biol. Chem. 270, 7829-7835, 1995

A:Title: A high affinity digoxin-binding protein displayed on M13 is functionally identical

A:Reference number: A56446; MUID:95229583; PMID:7713873

A:Accession: A56446

A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-268 <TAN>
A:Cross-references: UNIPARC:UPI000017C6D0; GB:U20617
C;Keywords: heterotetramer; immunoglobulin

Query Match 48.4%; Score 676.5; DB 2; Length 268;
Best Local Similarity 54.6%; Pred. No. 1.9e-40;
Matches 136; Conservative 36; Mismatches 68; Indels 9; Gaps 3;

Qy 1 MAQVLQVSGAEVKPGASVKVSKASGYTTSNYGITTWRQAPGQGLEWMGWI--NNGNT 58
Db 1 MAQVLQVSGAEVKPGASVKVSKASGYTTSNYGITTWRQAPGQGLEWMGWIAPANGIT 60

Qy 59 HYAQKFGQGVTTTDTSTSTAYMELRSDDTAVYICARDPRGDDPEYWGQGLTVTVSS 118
Db 61 KYDPKFGQKATIAADTSNTAYLQLSSLTSEDVAVYICASYLYLTRYENWGQGLTVTVSS 120

Qy 119 GGGSGGGGGGGGSEIVLTQSPSLPVLTLGPASISCRSSQNLVYSDGNTYLSWFOQRP 178
Db 121 GGGSGGGGGGGSDIELTQSPALMSASLGEKVTMSCRASSV-----NFIYVYQOKS 174

Qy 179 GQSPRLLYKVNRSQVDPDPFSGSGSGTDFTLKISRVEADIGVYVCMQGTHWPPRTFG 238
Db 175 DASPKLWYTYTSHLPVGPVAPRFGSGSGNSYSLTITSSMEGEDAATYVYCOQFTS--SPFTFG 233

Qy 239 QGTVKEIKR 247
Db 234 SGTKLEIKR 242

RESULT 3
JC5322
p53 specific single-chain antibody Pab421 - human
C;Species: Homo sapiens (man)
C;Date: 15-May-1997 #sequence_revision 15-May-1997 #text_change 18-Jul-1997
C;Accession: JC5322
R;Jannot, C.B.; Hynes, N.E.
Biochem. Biophys. Res. Commun. 230, 242-246, 1997
A;Title: Characterization of scFv-421, a single-chain antibody targeted to p53.
A;Reference number: JC5322; MUID:97168950; PMID:9016757
A;Accession: JC5322
A:Molecule type: mRNA
A:Residues: 1-233 <JAN>
A:Cross-references: UNIPARC:UPI000017C2DE
A;Experimental source: hybridoma cell
C;Comment: This protein specifically binds the tumor suppressor protein p53. It restores

Query Match 47.2%; Score 660; DB 2; Length 233;
Best Local Similarity 58.3%; Pred. No. 2.3e-39;
Matches 140; Conservative 33; Mismatches 57; Indels 10; Gaps 6;

Qy 8 QSGAEVKKPGASVKVSKASGYTTSNYGITTWRQAPGQGLEWMGIN--NGNTHYAKRFQ 65
Db 2 ESGAEVRSASVKLSCTTSGFNINDYTMHWYKKEPEQGLEWIGRIDPENGADMTRSSG 61

Qy 66 GRVTMTTSTSTAYMELRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGGGGSGG 125
Db 62 VKATMTADTSNTAYLQLSSLTSEDVAVYIC---NAGMD--YWGQGLTVTVSSGGGGSGG 116

Qy 126 GSGGGGSEIVLTQSPSLPVLTLGPASISCRSSQNLVYSDGNTYLSWFOQRPQSPRL 185
Db 117 RASGGGGSDIELTQSPASLAVSLGQATISCRASKS--VSTSGYSYMHWNQKRFQGPRL 175

Qy 186 IYKVNRSQVDPDRFGSGSGTDFTLKISRVEADIGVYVCMQGTHWPPRTFGQTKVEI 245
Db 176 IYLVNSLESGVAPRFGSGSGTDFTLNIHPVEEEDAATYVY--QHIRELTRSEG--GTLKEI 233

RESULT 4
B49002
Ig kappa chain V region, rheumatoid factor RF antibody - human (fragment)
C;Species: Homo sapiens (man)

C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: B49002
R;Stuber, F.; Lee, S.K.; Bridges Jr., S.L.; Koopman, W.J.; Schroeder, H.W.J.; Gaskin, F.;
Arthritis Rheum. 35, 900-904, 1992
A;Title: A rheumatoid factor from a normal individual encoded by VH2 and V kappa II gene
A;Reference number: A49002; MUID:92352481; PMID:1322670
A;Accession: B49002
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-114 <STU>
A:Cross-references: UNIPROT:Q9UL80; UNIPARC:UPI0000176B43
A;Experimental source: EBV-transformed lymphoblastoid cell line SSH23
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:110263, NCBIP:110264)
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;16-95/Domain: immunoglobulin homology <IMM>

Query Match 40.4%; Score 565; DB 2; Length 114;
Best Local Similarity 92.1%; Pred. No. 4.9e-33;
Matches 105; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 134 EIVLTQSPSLPVLTLGPASISCRSSQNLVYSDGNTYLSWFOQRPQSPRLLYKVSNRD 193
Db 1 DVVMTQSPSLPVLTLGPASISCRSSQSLVYSDGNTSLANWFOQRPQSPRLLYKVSNRD 60

Qy 194 SGVPDRFSGSGSGTDFTLKISRVEADIGVYVCMQGTHWPPRTFGQTKVEIKR 247
Db 61 SGVPDRFSGSGSGTDFTLKISRVEADIGVYVCMQGTHWPPPLTFGGGTVKEIKR 114

RESULT 5
S49572
Ig kappa chain precursor - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Mar-1995 #sequence_revision 14-Jul-1995 #text_change 21-Jan-2000
C;Accession: S49572
R;Giachino, C.; Padovan, E.; Lanzavecchia, A.
submitted to the EMBL Data Library, November 1994
A;Description: k+l+ dual receptor B cells are present in the human peripheral repertoire.
A;Reference number: S49571
A;Accession: S49572
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-114 <GIA>
A:Cross-references: UNIPARC:UPI0000116709; EMBL:Z46626; NID:G575261; PIDN:CAAB6596.1; PII
C;Superfamily: immunoglobulin V region; immunoglobulin homology
F;16-95/Domain: immunoglobulin homology <IMM>

Query Match 40.2%; Score 562; DB 2; Length 114;
Best Local Similarity 89.5%; Pred. No. 7.9e-33;
Matches 102; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

Qy 134 EIVLTQSPSLPVLTLGPASISCRSSQNLVYSDGNTYLSWFOQRPQSPRLLYKVSNRD 193
Db 1 DVVMTQSPSLPVLTLGPASISCRSSQSLVYDGTNTLANWFOQRPQSPRLLYKVSNRD 60

Qy 194 SGVPDRFSGSGSGTDFTLKISRVEADIGVYVCMQGTHWPPRTFGQTKVEIKR 247
Db 61 SGVPDRFSGSGSGTDFTLKISRVEADIGVYVCIQGTHWPPQYTFGGGTVKEIKR 114

RESULT 6
S23230
Ig kappa chain precursor V-J region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C;Accession: S23230
R;Kennedy, M.A.
J. Exp. Med. 173, 1033-1036, 1991
A;Title: Novel chromosome translocation caused by fusion of immunoglobulin heavy and light
A;Reference number: S23230; MUID:91178438; PMID:1840606
A;Accession: S23230

A;Title: Cloning of a gene encoding a lupus-associated human autoantibody V(K) region us
A;Reference number: S22902; MUID:91276289; PMID:1905262

A;Accession: S22902

A;Status: preliminary; translation not shown

A:Molecule type: mRNA

A;Residues: 1-142 <CHA>

A;Cross-references: UNIPARC:UIP0000176CAB; EMBL:X56510

C;Superfamily: immunoglobulin V region; immunoglobulin homology

K;Keywords: heterotetramer; immunoglobulin

F;47-126/Domain: immunoglobulin homology <IMM>

Query Match 39.2%; Score 547.5; DB 2; Length 142;
Best Local Similarity 87.9%; Pred.No.1e-31; Gaps 1;
Matches 102; Conservative 9; Mismatches 4; Indels 1;

Qy 130 GGSGEIVLTQSPLSPVTLTGQPASISCRSSQNLYSDGNTYLSWFOORPGSQSRRLIYKV 189
 |::|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 28 GSSEGVDMVTQSPLSLPVTLTGQPASISCRSSQSLVSHDGTNYLNWNFQRPGSQSRRLIYKV 87

Qy 190 SNRDSGVDPDRFGSGSGTDFTLKISRVEADIGVIYYCMOGTHWPPTFFGGTKVEIKR 245
 |||||||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 88 SNRDSGVDPDRFGSGSGTDFTLKISRVEADIGVIYYCMOGTHW-PFTFGGKTGLEI 142

RESULT 13

S42611

HUNK protein precursor - human

C;Species: Homo sapiens (man)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000

C;Accession: S42611

E;Spatz, L.A.; Williams, M.; Brender, B.; Desai, R.; Latov, N.
J. Neuroimmunol. 36, 29-39, 1992

A;Title: DNA sequence analysis and comparison of the variable heavy and light chain regions

A;Reference number: S42610; MUID:92138794; PMID:1370957

A;Accession: S42611

A>Status: preliminary

A:Molecule type: DNA

A;Residues: 1-133 <SPA>

A;Cross-references: UNIPARC:UIP000011378B; EMBL:X54137; PIDN:CAA38072.1; PIR:S42610

C;Superfamily: immunoglobulin V region; immunoglobulin homology

F;36-115/Domain: immunoglobulin homology <IMM>

Query Match 39.1%; Score 546.5; DB 2; Length 133;
Best Local Similarity 86.4%; Pred.No.1.e-31;
Matches 102; Conservative 8; Mismatches 7; Indels 1;

Qy 130 GGSGEIVLTQSPLSPVTLTGQPASISCRSSQNLYSDGNTYLSWFOORPGSQSRRLIYKV 189
 |::|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|~
Db 17 GSSEGVDMVTQSPLSPVTLTGQPASISCRSSQSLVSHDGTNYLNWNFQRPGSQSRRLIYKV 76

Qy 190 SNRDSGVDPDRFGSGSGTDFTLKISRVEADIGVIYYCMOGTHWPPTFFGGTKVEIKR 247
 |||||||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 77 SNRDSGVDPDRFGSGSGTDFTLKISRVEADIGVIYYCMOGAHW-PLTFGGGKTVEITR 133

RESULT 14

S40374

Ig kappa chain - human

C;Species: Homo sapiens (man)

C;Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 31-Dec-2004

C;Accession: S40374

R;Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993

A;Title: Expressed human immunoglobulin chi genes and their hypermutation.

A;Reference number: S40312; MUID:94080891; PMID:8258341

A;Accession: S40374

A>Status: preliminary; translation not shown

A:Molecule type: mRNA

A;Residues: 1-118 <KL&>

A;Cross-references: UNIPROT:Q8TCDO; UNIPARC:UIP0000176C9D; EMBL:X72484

C;Superfamily: immunoglobulin homology

K;Keywords: heterotetramer; immunoglobulin

F;20-99/Domain: immunoglobulin homology <IMM>

Search completed: August 9, 2006, 13:15:24
Job time : 41 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:06:26 ; Search time 300 Seconds
(without alignments)
804.764 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVQLVQSGAEVKKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	848.5	60.7	244	2	Q65ZC8_HUMAN
2	828	59.2	248	2	Q65ZQ7_MOUSE
3	764.5	54.7	255	2	Q6KB05_MOUSE
4	730.5	52.3	240	2	Q65ZC9_HUMAN
5	721.5	51.6	243	2	Q7TQM2_MOUSE
6	718.5	51.4	241	2	Q92IA6_MOUSE
7	717.5	51.3	487	2	Q65ZL2_MOUSE
8	664.5	47.5	218	2	Q925S1_MOUSE
9	565	40.4	114	2	Q9UL80_HUMAN
10	560.5	40.1	133	1	KV2F_HUMAN
11	557.5	39.9	239	2	Q8TCD0_HUMAN
12	493.5	35.3	239	2	Q6P491_HUMAN
13	493.5	35.3	500	2	Q6N091_HUMAN
14	477.5	34.2	469	2	Q7Z7P5_HUMAN
15	471.5	33.7	518	2	Q6N030_HUMAN
16	471	33.7	115	2	Q5F210_MOUSE
17	465.5	33.3	239	2	Q8NEK0_HUMAN
18	462	33.0	124	2	Q9UL92_HUMAN
19	461.5	33.0	125	2	Q9UL95_HUMAN
20	458.5	32.8	113	1	KV2G_MOUSE
21	458.5	32.8	119	2	Q9UL94_HUMAN
22	456.5	32.7	117	1	KV2E_HUMAN
23	455.5	32.6	239	2	Q58E08_MOUSE
24	451	32.3	240	2	Q6PIH6_HUMAN
25	450.5	32.2	113	1	KV2B_HUMAN
26	450	32.2	498	2	Q6N041_HUMAN
27	448.5	32.1	113	1	KV2D_HUMAN
28	446	31.9	170	2	Q925S2_MOUSE
29	445	31.8	500	2	Q9BRV0_HUMAN
30	444	31.8	159	2	Q96Q80_HUMAN
31	438.5	31.4	219	2	Q65ZC0_MOUSE

32	433.5	31.0	480	2	Q6P089_HUMAN	Q6P089 homo sapien
33	428	30.6	86	2	Q7Z3V5_HUMAN	Q7Z3V5 homo sapien
34	426	30.5	115	1	KV2A_HUMAN	P01614 homo sapien
35	425.5	30.4	119	2	Q9GYZ2_MOUSE	Q9GYZ2 mus musculus
36	425	30.4	147	1	HVIC_HUMAN	P01744 homo sapien
37	422	30.2	112	1	KV2C_HUMAN	P01616 homo sapien
38	421	30.1	117	1	HVIB_HUMAN	P01742 homo sapien
39	419.5	30.0	458	2	Q5BJZ2_RAT	Q5BJZ2 rattus norv
40	417.5	29.9	112	2	Q53VP8_MOUSE	Q53VP8 mus musculus
41	417	29.8	116	2	Q9UL89_HUMAN	Q9UL89 homo sapien
42	415.5	29.7	475	2	Q6N095_HUMAN	Q6N095 homo sapien
43	414.5	29.6	497	2	Q8WY24_HUMAN	Q8WY24 homo sapien
44	414	29.6	519	2	Q5BEM2_HUMAN	Q5BEM2 homo sapien
45	410.5	29.4	234	2	Q5XKG4_MOUSE	Q5XKG4 mus musculus

ALIGNMENTS

RESULT 1

Q65ZC8_HUMAN PRELIMINARY; PRT; 244 AA.
AC Q65ZC8; integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 11-OCT-2004, entry version 9.
DE Single-chain Fv (Fragment).
GN Name=scFv;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97362799; PubMed=9219263; DOI=10.1038/nbr0797-629;
RA Kontermann R.E., Wing M.G., Winter G.;
RT "Complement recruitment using bispecific diabodies."
RL Nat. Biotechnol. 15:629-631(1997).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; Y13057; CAA73500.1; -; mRNA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_v.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 2.
DR SMART; SM00406; IG; 2.
DR PROSITE; PS50835; IG_LIKE; 2.
KW Immunoglobulin domain.
FT NON_TER 1
FT NON_TER 244
SQ SEQUENCE 244 AA; 26127 MW; 4B1F17868338F2BF CRC64;

Query Match 60.7%; Score 848.5; DB 2; Length 244;

Best Local Similarity 65.6%; Pred. No. 6.1e-61;

Matches 164; Conservative 34; Mismatches 41; Indels 11; Gaps 4;

QY 3 QVQLVQSGAEVKKPGASVKVKSCASGYTFSTNYGITVWRQAPGGQLEMMGW---NNGNTHY 60

1 QVQLVQSGAEVKKPGDSVKVKSCASGYTFSDHYMHVWRQAPGGQLEMMGWIDPNNGDTRF 60

61 AQKFGQRTVTMTDTSTAYMELSLRSDDTAVVYCARPRGD---DEPYWGQGITLVTS 117

61 AQRFGQRTVTMTDTSTAYMEVSLRSDDTAVVYCAREGTGSATYGMVWGQGITLVTS 120

118 SGGGSGGGSGGGSGSEIVLTQSLPVLTPGAPASISCRSSQNLVYSDGNTVLSNFFQQR 177

121 SGGGSGGGSGGGSGGGSDIQMTQSPSLASISIDRVITTCRASEGIYH-----WLAWTQQK 175

178 PQSPRRLIYKVSNRDSDGVPDRFSGSGSGTDTFTLKISRVEAEDIGVYVCMQGTWPPRTF 237

```
Db 176 PGKAPKFLIYKASLASGAPRFGSGSGTDTLTITSSLPQDDFATYYCQOYSNY-PLTF 234
Qy 238 GGGTKVEIKR 247
Db 235 GGGTKLEIKR 244

RESULT 2
Q65ZQ7_9MURI
ID Q65ZQ7_9MURI PRELIMINARY; PRT; 248 AA.
AC Q65ZQ7;
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE B3(Fv)-PE40 (Fragment).
GN Name=B3(Fv)-PE40;
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92020904; PubMed=1924323;
RA Brinkmann U., Pai L.H., FitzGerald D.J., Willingham M., Pastan I.;
RT "B3(Fv)-PE38DEL, a single-chain immunotoxin that causes complete
RT regression of a human carcinoma in mice.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:8616-8620 (1991).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DE EMBL; S57990; AAB19971.2; -; mRNA.
DR SMR; Q65ZQ7; 4-247.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_V.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 2.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PS50835; IG LIKE; 2.
KW Immunoglobulin domain.
FT NON_TER 248 248
SQ SEQUENCE 248 AA; 26634 MW; 7A3759B43B570950 CRC64;

Query Match 59.2%; Score 828; DB 2; Length 248;
Best Local Similarity 64.2%; Pred. No. 2.9e-59;
Matches 158; Conservative 41; Mismatches 43; Indels 4; Gaps 3;

Qy 4 VOLVQSGAEVKKPGASVKASCKASYTFSTNYGIVTWVROAPGGLEWMGNNGNTH--YA 61
Db 3 VKLVESGGGLVQPGGSLKSLCATSGFTFSDYYWYWRQTPKRLSEWYVAINDDSSAAYS 62
Qy 62 QKQGRVMTTDTSTSTAYMELRSDDTAVYYCARD-PRGDDDFYWGQGLTVTVSSGG 120
Db 63 DTVKGRFTISRDNARTLYLQMSRLKSEDTALYSCARGLANGAWFAYWGQGLTVTVSSG 122
Qy 121 GSGGGGGGGGGSEIVLTQSPSLPVTTLGQPASISCRSSQNLVYSDGNTLYSWFQPRGQ 180
Db 123 GSGGGGGGGGGSDVLTQSPSLPVSGLDQASISCRSSQIIIVHSGNTLYEWYLQKPGQ 182
Qy 181 SPRLIYKVNRSQVPRFSGSGGTDTFTLKISRVEAEDIGVYCMQGTTHWPPRFGQ 240
Db 183 SPKLLIYKVNRSQVPRFSGSGGTDTFTLKISRVEAEDIGVYCMQGTTHWPPRFGQ 241
Qy 241 TKVEIK 246
Db 242 TKLEIK 247

RESULT 3
Q6KB05_MOUSE
```

```
ID Q6KB05_MOUSE PRELIMINARY; PRT; 255 AA.
AC Q6KB05;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE SCFV B8E5 protein (Fragment).
GN Name=scfv B8E5;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=Balb/c;
RX PubMed=15485827; DOI=10.1074/jbc.M407213200;
RA Peter J.C., Wallukat G., Tugler J., Maurice D., Roegel J.C.,
RA Briand J.P., Hoebeke J.;
RT "Modulation of the M2 muscarinic acetylcholine receptor activity with
RT monoclonal anti-M2 receptor antibody fragments.";
RL J. Biol. Chem. 279:55697-55706 (2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DE EMBL; AJ746180; CAG34081.1; -; Other_DNA.
DR HSSP; P01837; 1KCR.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_V.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 2.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PS50835; IG LIKE; 2.
KW Immunoglobulin domain.
FT NON_TER 255 255
SQ SEQUENCE 255 AA; 27445 MW; B68BD38395DF713B CRC64;

Query Match 54.7%; Score 764.5; DB 2; Length 255;
Best Local Similarity 59.6%; Pred. No. 4.5e-54;
Matches 149; Conservative 42; Mismatches 52; Indels 7; Gaps 5;

Qy 3 QVQLVQSGAEVKKPGASVKASCKASYTFSTNYGIVTWVROAPGGLEWMGNNGN--THY 60
Db 1 QVQLQQSGDLVKPGGSLKVSAAAGFTFSSYGMVWRQTPDKRLSEWVAITISGGSYTY 60
Qy 61 AQKQGRVMTTDTSTSTAYMELRSDDTAVYYCAR--DPRGDDDE-PYWGQGLTVTVS 117
Db 61 PDSVKGRTISRDNARTLYLQMSRLKSEDTAMYYCARHINYDGAFDYWGQGLTVTVS 120
Qy 118 GSGGGGGGGGGSEIVLTQSPSLPVTTLGQPASISCRSSQNLVYS-DGNTYLSWFQQ 176
Db 121 GSGGGGGGGGGGGSDIVNAQSPSSLSVAGKIVMSCKSSQSLNSRNKQKYLAWYQQ 180
Qy 177 RFQGSFRLIYKVNRSQVPRFSGSGGTDTFTLKISRVEAEDIGVYCMQGTTHWPPRT 236
Db 181 KFGQSPKLLIYKASTRESQVPRFSGSGGTDTFTLTISVQAEDLAVYYC-QNDHSYPLT 239
Qy 237 FCGQTKVEIK 246
Db 240 FGAGTKLEIK 249

RESULT 4
Q65ZC9_HUMAN
ID Q65ZC9_HUMAN PRELIMINARY; PRT; 240 AA.
AC Q65ZC9;
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Single-chain Fv (Fragment).
GN Name=scFv;
OS Homo sapiens (Human).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP STRAIN=C1G/7;
 RC NUCLEOTIDE SEQUENCE.
 RX MEDLINE=97362799; PubMed=9219263; DOI=10.1038/nbt0797-629;
 RA Kontermann R.E., Wing M.G., Winter G.;
 RT "Complement recruitment using bispecific diabodies.";
 RL Nat. Biotechnol. 15:629-631(1997).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; Y13056; CAA73499.1; -, mRNA.
 DR LinkHub; Q652C9; -.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003596; IG_v.
 DR InterPro; IPR013106; V-set.
 DR SMART; SM00409; IG; 2.
 DR SMART; SM00406; IGV; 2.
 DR PROSITE; PS00835; IG LIKE; 2.
 KW Immunoglobulin domain.
 FT NON_TER 240 240
 FT SEQUENCE 240 AA; 25569 MW; FDCFD3645F64B373 CRC64;
 SQ
 Query Match 52.3%; Score 730.5; DB 2; Length 240;
 Best Local Similarity 56.3%; Pred. No. 2.5e-51;
 Matches 139; Conservative 47; Mismatches 52; Indels 9; Gaps 4;
 QY 3 QVQLVQSGAEVKKPKASVKVCKASGYTFSNVTGIVTWRQAPQGQLEWGMWGI--NNGNTHY 60
 DB 1 QVQLVQSGGLVQPGGSLRLCSAASGFTFSYGMHWVRQAPKGLEWVAVISYDSNKYY 60
 QY 61 AQKPGQRTVMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPYWGQGTTLTVSSGG 120
 DB 61 ADSVKGRTISRDNSKNTLYLQMSLRAEDTAVYYCARD--WGDSLDPWGKGLTVTVSSGG 119
 QY 121 GSGGGSGGGGGSEIVLTQSPSLPTVLTGPASISCRSQNLVYSDGNTYLSWFOQRPGQ 180
 DB 120 GSGGGSGGGGGSDIQMTQSPSTLSASIGDRVTITCRASEGIY-----RWLAWYQKPKG 174
 QY 181 SPRLIYKVNDRSGVDPDRFSGSGGTDTLTKISRVEAEDIGVYCMQGTWPPPTFGQ 240
 DB 175 APKLLIYKASSLASRAPRFSFGSGGTDTLTLSLQPDPTATYYCQYSNY-PLTFGGG 233
 QY 241 TKVEIKR 247
 DB 234 TKLEIKR 240
 RESULT 5
 Q7TQM2 MOUSE
 ID Q7TQM2 MOUSE PRELIMINARY; PRT; 243 AA.
 AC Q7TQM2
 DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
 DT 01-OCT-2003, sequence version 1.
 DT 07-FEB-2006, entry version 15.
 DE ScFv 6H8 protein (Fragment).
 GN Name=Igk-V33; Synonyms=scFv 6H8;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Balb/c;
 RX MEDLINE=22853226; PubMed=12860977; DOI=10.1074/jbc.M306877200;
 RA Peter J.C., Eftekhari P., Billiald P., Wallukat G., Hoebeke J.;

RT "scFv single chain antibody variable fragment as inverse agonist for
 the beta-2 adrenergic receptor.";
 RL J. Biol. Chem. 278:36740-36747(2003).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; AJ574851; CAE00495.1; -, Genomic_DNA.
 DR HSSP; P01751; 1A6W.
 DR SMR; Q7TQM2; 1-236.
 DR MGI; MGI:96514; IGK-V33.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003596; IG_v.
 DR InterPro; IPR013106; V-set.
 DR SMART; SM00409; IG; 2.
 DR SMART; SM00406; IGV; 2.
 DR PROSITE; PS00835; IG LIKE; 2.
 KW Immunoglobulin domain.
 FT NON_TER 243 243
 FT SEQUENCE 243 AA; 25976 MW; BEFF64D2DCF4F76 CRC64;
 SQ
 Query Match 51.6%; Score 721.5; DB 2; Length 243;
 Best Local Similarity 56.9%; Pred. No. 1.4e-50;
 Matches 140; Conservative 43; Mismatches 52; Indels 11; Gaps 4;
 QY 3 QVQLVQSGAEVKKPKASVKVCKASGYTFSNVTGIVTWRQAPQGQLEWGMWGI--NNGNTHY 60
 DB 1 QVQLVQSGGLVQPGGSLRLCSAASGFTFSYGMHWVRQAPKGLEWVAVISYDSNKYY 60
 QY 61 AQKPGQRTVMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPYWGQGTTLTVSSGG 120
 DB 61 DEKFNKGLITVDTSSSTAYMHLSSLASDASVYYCARGGRGLD--VWGAGTTLTVSSGG 118
 QY 121 GSGGGSGGGGGSEIVLTQSPSLPTVLTGPASISCRSQNLVYSDGNTYLSWFOQRPGQ 180
 DB 119 GSGGGSGGGGGSDIQMTQSPSTLSASIGDRVTITCRASEDIY-----NRLAWYQKPKG 173
 QY 181 SPRLIYKVNDRSGVDPDRFSGSGGTDTLTKISRVEAEDIGVYCMQGTWPPPTFGQ 240
 DB 174 APRLLISGNTSLTGVDPFRFSFGSGGKDYLTSLTQEDVATYYCQ--YWSRTTGGG 231
 QY 241 TKVEIK 246
 DB 232 TKLEIK 237
 RESULT 6
 Q921A6 MOUSE
 ID Q921A6 MOUSE PRELIMINARY; PRT; 241 AA.
 AC Q921A6
 DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
 DT 01-DEC-2001, sequence version 1.
 DT 07-FEB-2006, entry version 19.
 DE Anti-CEA 79 single chain Fv (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=98170165; PubMed=9509426;
 RA Chung J.H., Choi S.J., Kim H.J., Kim I.J., Choi I.H., Lee S.D.,
 Yi K.S., Suh P.G., Ryu S.H., Chung H.K.;
 RT "Cloning and characterization of cDNAs encoding VH and VL of a
 monoclonal anti-CEA antibody (CEA 79) cross-reactive with NCA-95 and
 generation of a single-chain Fv molecule (scFv).";
 RL Mol. Cells 7:816-819(1997).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=91341421; PubMed=1908510; DOI=10.1084/jem.174.3.613;
 RA Stark S.E., Caton A.J.;

RT "Antibodies that are specific for a single amino acid interchange in a
RL protein epitope use structurally distinct variable regions.";
J. Exp. Med. 174:613-624(1991).

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----

DR EMBL; U88067; AAB48044.1; -; mRNA.

DR PIR; S19965; S19965.

DR PIR; S19967; S19967.

DR PIR; S19968; S19968.

DR PIR; S26325; S26325.

DR HSP; P01607; 1BWW.

DR InterPro; IPR003599; IG.

DR InterPro; IPR007110; IG-like.

DR InterPro; IPR003598; IG C2.

DR InterPro; IPR003596; IG V.

DR InterPro; IPR013106; V-set.

DR SMART; SM00409; IG; 2.

DR SMART; SM00406; IGV; 2.

DR PROSITE; PS00835; IG LIKE; 2.

KW Immunoglobulin domain.

FT NON TER 1 241

FT NON TER 241 241

SQ SEQUENCE 241 AA; 26086 MW; 0276887248B9C771 CRC64;

Query Match 51.4%; Score 718.5; DB 2; Length 241;

Best Local Similarity 54.8%; Pred. No. 2.4e-50;

Matches 138; Conservative 42; Mismatches 53; Indels 19; Gaps 6;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGQGLEWMGWIN--NGNTHY 60

DB 1 QVQLQSGPELKPGETVKISKASGYTFSTNYGIVTWRQAPGQGLEWMGWINTYGEPTY 60

QY 61 AQKFGQVMTTDTSTAYMELSLRSDDTAVVYCARDPRGDDPEYVQGGTLVTVSSGG 120

DB 61 ADPKGAPAFSLTASATAYIQINLNKEDTATVFCARQDLRLRYDYVQGGTLVTVSSGG 120

QY 121 GSGSGSGSGGGSEIVLTQSPSLPVTIGQPASISCRSSQNLVSDGNTYLSWFOORPGQ 180

DB 121 GSGSGSGSGGGSEIVLTQSPSLPVTIGQPASISCRSSQNLVSDGNTYLSWFOORPGQ 180

DB 121 GSGSGSGSGGGSEIVLTQSPSLPVTIGQPASISCRSSQNLVSDGNTYLSWFOORPGQ 175

QY 181 SPRL-----IYKVNDRSGVDPDRFSGSGGTDFTLKISRVEAEDIGVYCMQGTTHWPP-R 235

DB 176 GPRSAHTLHY-----IQGIPSRFSGSGGRDYSFISNLPEDATYCYL---HYDNLH 228

QY 236 TFGQGTKEIKR 247

DB 229 TFGGTTKLELR 240

RESULT 7

Q65ZL2_9MURI

ID Q65ZL2_9MURI PRELIMINARY; PRT; 487 AA.

AC Q65ZL2;

DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.

DT 11-OCT-2004, sequence version 1.

DT 07-FEB-2006, entry version 13.

DE FV/M4.

GN Name=M4-IPN-<tau>;

OS Mus sp.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Murinae; Mus.

OX NCBI_TaxID=10095;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=96272580; PubMed=8688499;

RA Qi Y., Xiang J.;

RT "A genetically engineered single-gene-encoded anti-TAG72 chimeric

antibody secreted from myeloma cells.";

Hum. Antibodies Hybridomas 6:161-166(1995).

CC -----

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----

DR EMBL; S82493; AAB37424.2; -; mRNA.

DR InterPro; IPR003599; IG.

DR InterPro; IPR007110; IG-like.

DR InterPro; IPR003597; IG C1.

DR InterPro; IPR003598; IG C2.

DR InterPro; IPR003006; IG MHC.

DR InterPro; IPR003596; IG V.

DR Pfam; PF07654; C1-set; 2.

DR SMART; SM00409; IG; 2.

DR SMART; SM00407; IGV; 1.

DR SMART; SM00406; IGV; 2.

DR PROSITE; PS00835; IG LIKE; 4.

DR PROSITE; PS00290; IG MHC; 1.

SQ SEQUENCE 487 AA; 53578 MW; C7BAB69F30555504 CRC64;

Query Match 51.3%; Score 717.5; DB 2; Length 487;

Best Local Similarity 56.6%; Pred. No. 6.5e-50;

Matches 141; Conservative 46; Mismatches 45; Indels 17; Gaps 7;

QY 2 AQVLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGQGLEWMGWINGN--TH 59

DB 19 SQVQLQSDAEIVKPGASVKISKASGYTFTHAIHWAKQKPGQGLEWIGYISPGNDIK 78

QY 60 YAQKFGQVMTTDTSTAYMELSLRSDDTAVVYCARDPRGDDPEYVQGGTLVTVSSG 119

DB 79 YNEKFKGRATLTADKSSSTAYMQLNSLTSEDSAVYFCKSYTG---HWQGGTLT-GSG 133

QY 120 GSGSGSGSGGGSEIVLTQSPSLPVTIGQPASISCRSSQNLVSDGNTYLSWFOORPG 179

DB 134 GSGSGSGSGGGSEIVLTQSPSLPVTIGQPASISCRSSQNLVSDGNTYLSWFOORPG 188

QY 180 QSPRLIYKVNDRSGVDPDRFSGSGGTDFTLKISRVEAEDIGVYCMQ---GTHWPPRTF 237

DB 189 KSPQLVVAATNLADGVPSRFSGSGGTQYSLKNSLQSDFGSYVQHPWGT---PYTF 245

QY 238 GQGTKEIKR 246

DB 246 GGGTKEIKR 254

RESULT 8

Q925S1_MOUSE

ID Q925S1_MOUSE PRELIMINARY; PRT; 218 AA.

AC Q925S1;

DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.

DT 01-DEC-2001, sequence version 1.

DT 07-FEB-2006, entry version 18.

DE MRP5 (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=BA1B/c;

RX PubMed=11819679;

RA Cui D.X., Zeng G.Y., Wang F., Xu J.R., Ren D.Q., Guo Y.H., Tian F.R.,

Yan X.J., Hou Y., Su C.Z.;

RT "Mechanism of exogenous nucleic acids and their precursors improving

the repair of intestinal epithelium after gamma-irradiation in mice.";

World J. Gastroenterol. 6:709-717(2000).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=BA1B/c;

RA Cui D., Zeng G., Yan X., Li X., Su C.;

RT "Cloning of mouse genes related to repairing of intestinal epithelium

of the irradiated mice by treatment with the intestinal RNA of mice of

the same strain.";

RT

```

RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF240168; AAK43733.1; -; mRNA.
DR HSSP; P01665; 1QNZ.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG.v.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin domain.
FT NON TER 218
SQ SEQUENCE 218 AA; 23013 MW; 527E4FA8F7982817 CRC64;

Query Match 47.5%; Score 664.5; DB 2; Length 218;
Best Local Similarity 57.5%; Pred. No. 5.3e-46;
Matches 126; Conservative 42; Mismatches 48; Indels 3; Gaps 2;

QY 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQLEWGMWIN--NGNT 58
DB 1 MAQVQLQSGPELKPKGETWRISCKASGYTFTAGQWQKMPKGLKWIWINTHSGVP 60

QY 59 HYAQKFGQRTMTTDTSTAYMELRLSRDDTAVYVCARDPRGDEPYWGQGLTVTVSS 118
DB 61 KYAEFGKGRFAFSLSTASTAYLQISNLKNDTATYFCWRWDYDGGPAYWGQGLTVTVSS 120

QY 119 GGGSGGGGGGGSEIVLTQSPISLPLVTLPQASISCRSSQNLVYSDGNTYLSWFOQRP 178
DB 121 GGGSGGGGGGGSDIVLTQSPASLASVLSQRTATISCRASES-VDNIGISFMNWFQKRP 179

QY 179 GQSPRLIYKVNEDSGVDFRFGSGSGTDTLTKISRVE 217
DB 180 GQPPKLLIYAASQGGVPGAGLIASGSGTDFSLNIYPME 218

RESULT 9
Q9UL80 HUMAN PRELIMINARY; PRT; 114 AA.
AC Q9UL80;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Myosin-reactive immunoglobulin light chain variable region (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98271139; PubMed=9614934; DOI=10.1006/clin.1998.4531;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92352481; PubMed=1322670;
RA Stuber F., Lee S.K., Bridges S.L. Jr, Koopman W.J., Schroeder H.W. Jr,
RA Gaskin F., Fu S.M.;
RT "A rheumatoid factor from a normal individual encoded by VH2 and V
RT kappa II gene segments.";
RL Arthritis Rheum. 35:900-904(1992).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93170387; PubMed=8436174;
RA Wagner S.D., Luzzatto L.;
RT "V kappa gene segments rearranged in chronic lymphocytic leukemia are

```

```

RT distributed over a large portion of the V kappa locus and do not show
RT somatic mutation.";
RL Eur. J. Immunol. 23:391-397(1993).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92289816; PubMed=1601042;
RA Huber C., Klobeck H.G., Zachau H.G.;
RT "Ongoing V kappa-J kappa recombination after formation of a productive
RT V kappa-J kappa coding joint.";
RL Eur. J. Immunol. 22:1561-1565(1992).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF035034; AAD56270.1; -; mRNA.
DR PIR; B49002; B49002.
DR PIR; S23638; S23638.
DR PIR; S34094; S34094.
DR PIR; S34095; S34095.
DR HSSP; P01625; ILVE.
DR SMR; Q9UL80; 1-114.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG.v.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin domain.
FT NON TER 114
SQ SEQUENCE 114 AA; 12775 MW; 070E31E210D1CB01 CRC64;

Query Match 40.4%; Score 565; DB 2; Length 114;
Best Local Similarity 92.1%; Pred. No. 3.2e-38;
Matches 105; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 134 EIVLTQSPISLPLVTLPQASISCRSSQNLVYSDGNTYLSWFOQRPQSPRLIYKVSNRD 193
DB 1 DVVMTQSPISLPLVTLPQASISCRSSQSPVSDGNTYLSWFOQRPQSPRLIYKVSNRD 60

QY 194 SGVPRFRFGSGSGTDTLTKISRVEADIGVYCMQGTHTWPPRTFGQTKVEIKR 247
DB 61 SGVPRFRFGSGSGTDTLTKISRVEADVGYYCMQGTHTWPPRTFGQTKVEIKR 114

RESULT 10
KV2F HUMAN STANDARD; PRT; 133 AA.
AC P06310;
DT 01-JAN-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-JAN-1988, sequence version 1.
DT 07-MAR-2006, entry version 41.
DE IG kappa chain V-II region RPMI 6410 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=86041852; PubMed=2997711;
RA Klobeck H.G., Meindl A., Combratio G., Solomon A., Zachau H.G.;
RT "Human immunoglobulin kappa light chain genes of subgroups II and
RT III.";
RL Nucleic Acids Res. 13:6499-6513(1985).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; Z00020; CAA77315.1; -; Genomic DNA.
DR PIR; A01890; K2HURP.
DR HSSP; Q99M37; 1I91.

```

DR SMR; P06310; 21-133.
 DR Ensembl; ENSG00000173758; Homo sapiens.
 DR LinkHub; P06310; -
 DR GO; GO:0005576; C:extracellular region; NAS.
 DR GO; GO:0003823; F:antigen binding; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR InterPro; IPR013106; V-set.
 DR Pfam; PF07686; V-set; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00406; IGv; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 KW Immunoglobulin domain; Immunoglobulin V region; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 133
 FT REGION 21 43
 FT REGION 44 59
 FT REGION 60 74
 FT REGION 75 81
 FT REGION 82 113
 FT REGION 114 122
 FT REGION 123 132
 FT DISULFID 43 113
 FT NON TER 133 133
 SQ SEQUENCE 133 AA; 14707 MW; 513CCAF3673009EE CRC64;
 Query Match 40.1%; Score 560.5; DB 1; Length 133;
 Best Local Similarity 89.8%; Pred. No. 8.8e-38;
 Matches 106; Conservative 6; Mismatches 5; Indels 1; Gaps 1;
 QY 130 GGGSEIVLTQSPVLTGQPASTISCRSSQNLVYSGNTYLSWFFQRPQGSPRLIYKV 189
 DB 17 GSGSDVMTQSPVLTGQPASTISCRSSQSLVYSGNTYLSWFFQRPQGSPRLIYKV 76
 QY 190 SNRDGVPDRFSGSGGTDFTLTKISRVAEIDGVYCMQGHWPRTFGQGTQVEIKR 247
 DB 77 SNRDGVPDRFSGSGGTDFTLTKISRVAEIDGVYCMQGHWSWTFQGTQVEIKR 133
 RESULT 11
 Q8TCD0 HUMAN
 ID Q8TCD0 HUMAN PRELIMINARY; PRT; 239 AA.
 AC Q8TCD0;
 DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
 DT 01-JUN-2002, sequence version 1.
 DT 07-FEB-2006, entry version 24.
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Lung;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heiton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Shailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Lung;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92285150; PubMed=1598223;
 RA Hirabayashi Y., Munakata Y., Sasaki T., Sano H.;
 RT "Variable regions of a human anti-DNA antibody O-81 possessing lupus
 nephritis-associated idiotype.";
 RL Nucleic Acids Res. 20:2601-0(1992).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92201291; PubMed=1551402;
 RA Lautner-Rieske A., Huber C., Meindl A., Pargent W., Schable K.F.,
 Thiebe R., Zocher I., Zachau H.G.;
 RT "The human immunoglobulin kappa locus. Characterization of the
 duplicated A regions.";
 RL Eur. J. Immunol. 22:1023-1029(1992).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=94080891; PubMed=8258341;
 RA Klein R., Jaenichen R., Zachau H.G.;
 RT "Expressed human immunoglobulin kappa genes and their hypermutation.";
 RL Eur. J. Immunol. 23:3248-3262(1993).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=93170387; PubMed=8436174;
 RA Wagner S.D., Luzzatto L.;
 RT "v kappa gene segments rearranged in chronic lymphocytic leukemia are
 distributed over a large portion of the v kappa locus and do not show
 somatic mutation.";
 RL Eur. J. Immunol. 23:391-397(1993).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 CC EMBL; BC022362; AAH22362.1; -; mRNA.
 DR PIR; S22658; S22658.
 DR PIR; S34095; S34095.
 DR PIR; S40324; S40324.
 DR PIR; S40374; S40374.
 DR PIR; S42267; S42267.
 DR PIR; S42268; S42268.
 DR HSP; P01834; I172.
 DR SMR; Q8TCD0; 21-237.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003597; Ig_c1.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 DR InterPro; IPR013106; V-set.
 DR Pfam; PF07654; C1-set; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00407; IGc1; 1.
 DR SMART; SM00406; IGv; 1.
 DR PROSITE; PS00835; IG_LIKE; 2.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 239 AA; 26235 MW; FACBDC3A3B0387ID CRC64;
 Query Match 39.9%; Score 557.5; DB 2; Length 239;
 Best Local Similarity 87.3%; Pred. No. 3.1e-37;
 Matches 103; Conservative 10; Mismatches 4; Indels 1; Gaps 1;
 QY 130 GGGSEIVLTQSPVLTGQPASTISCRSSQNLVYSGNTYLSWFFQRPQGSPRLIYKV 189

[illegible]

Search completed: August 9, 2006, 13:14:39
Job time : 303 secs